DOCUMENT RESUME

ED 231 672 SE 042 214

AUTHOR	Blosser, Patricia E., Ed.; Mayer, Victor J., Ed.
TITLE	Investigations in Science Education. Vol. 9, No. 2.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	83
NOTE	76p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $8.00, $2.25 single copy).
PUB TYPE	Collected Works - Serials (022) -- Information Analyses - ERIC Information Analysis Products (071) -- Guides - Non-Classroom Use (055)
JOURNAL CIT	Investigations in Science Education; v9 n2 1983
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Cognitive Development; *Cognitive Tests; Elementary Secondary Education; Higher Education; Inservice Teacher Education; *Learning; Preservice Teacher Education; *Science Education; Science Instruction; Science Teachers; Secondary School Science; Student Attitudes; *Teacher Attitudes
IDENTIFIERS	Marine Education; Meta Analysis; *Science Education Research

ABSTRACT

Critiques of 10 articles reporting research studies focused on various aspects of attitude development in science education and of six research articles related to investigations of learning and cognitive development are contained in this issue. Seven of the attitude articles describe investigations designed to change the attitudes of both pre-service and in-service elementary school teachers so that these individuals will hold more positive attitudes toward the teaching of science to elementary school pupils. Two other attitude articles are focused on students' reactions to the use of live animals in high school biology courses. The tenth attitude article provides a discussion of students' attitudes about marine education. Within the cluster of six articles about learning and cognitive development in science are included research focused on (1) correlates of formal reasoning, (2) the development of seven Science Reasoning Tasks for use in assessing the cognitive development of individuals when tested in groups of 20 or more, (3) formal operational ability and the teaching of science processes, (4) effects of textbook study questions on student learning in science, (5) synthesizing research on ability and science learning, and (6) synthesizing the effects of age and developmental levels on science learning. (PEB)

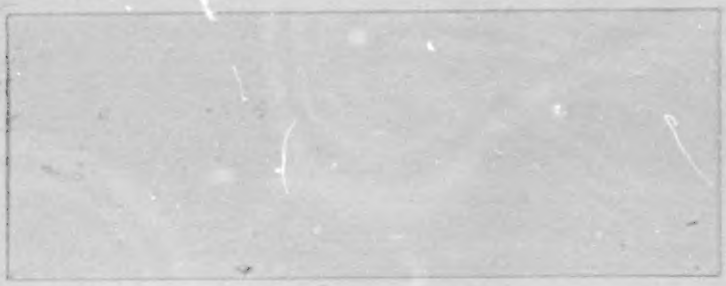
****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED231672

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SCIENCE EDUCATION INFORMATION REPORT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Robert W. Howe
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE
in cooperation with
Center for Science and Mathematics Education
The Ohio State University

SE042214

INVESTIGATIONS IN SCIENCE EDUCATION

Editor

Patricia E. Blosser
The Ohio State University

Associate Editor

Victor J. Mayer
The Ohio State University

Advisory Board

Robert L. Steiner (1983)
University of Puget Sound

Gerald Neufeld (1984)
Brandon University

Anton E. Lawson (1984)
Arizona State University

Lowell J. Bethel (1985)
University of Texas

Rodger Bybee (1984)
Carlton College

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 9, Number 2, 1983

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, OH 43210

Subscription Price: $8.00 per year. Single Copy Price: $2.25
Add $1.00 for Canadian mailings and $3.00 for foreign mailings.

Volume 9, Number 2, 1983

NOTES FROM THE EDITOR iii

ATTITUDES 1

Piper, Martha K. "The Investigation of Attitude Changes in Elementary Preservice Teachers in a Competency-Based, Field-Oriented Science Methods Course," Piper, M. K. and K. D. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by MICHAEL J. PADILLA. 3

Tamir, Pinchas and Efrat Sever. "Students' Attitudes Towards the Use of Animals in Biology Teaching," The American Biology Teacher, 42(2): 100-103, 1980.
Abstracted by PAUL JOSLIN 7

Piper, Martha K. and Kenneth D. Moore. "The Effect of a Physics Course for Elementary Teachers on Attitudes Toward Science of Preservice Elementary Teachers," in Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by WARREN J. SAFTER and JAMES R. OKEY. 11

Moore, K. D. and M. K. Piper. "Factors Underlying Student Teachers' Attitudes Toward Science in a Preservice Elementary Program," Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by KENNETH G. TOBIN. 15

Krockover, Gerald H. and Harold H. Jaus. "A Comparative Analysis of Teacher Attitudes in Two National Science Foundation Implementation Programs." In M. Piper and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by KENNETH S. RICKER 18

Piper, Martha K. and Linda Hough. "Attitude Changes of Classroom Teachers Cooperating with the Field Component of an Elementary Science Methods Course." In Piper, M. and K. Moore, Eds., Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by JOSEPH P. RILEY 21

Silberstein, Moshe and Pinchas Tamir. "Factors Which Affect Students' Attitudes Towards the Use of Living Animals in Learning Biology," Science Education, 65(2): 119-130, 1981.
Abstracted by CONSTANCE M. Perry. 24

Moore, Kenneth D. "An Investigation of the Science Background of Elementary Teachers," Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by JOSEPH P. RILEY 32

Piper, Martha K. "Comparison of Attitude Changes of Preservice Elementary Teachers Toward Science Between a Traditional Science Methods Course and a Non-Traditional Science Methods Course," in Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.
Abstracted by LEON W. BENEFIELD, JR. and JAMES R. OKEY. 35

Fortner, R. W. and T. G. Teates. "Baseline Studies for Marine Education: Experiences Related to Marine Knowledge and Attitudes." Journal of Environmental Education, 11(4): 11-19, 1980.
Abstracted by RICHARD M. SCHLENKER. 39

Response by Fortner to Schlenker's Critique 44

LEARNING, COGNITIVE DEVELOPMENT. 45

Linn, Marcia C., Steven Pulos, and Adrienne Gans. "Correlates of Formal Reasoning: Content and Problem Effects." Journal of Research in Science Teaching, 18(5): 435-447, 1981.
Abstracted by ANTON E. LAWSON 47

Shayer, Michael, Philip Adey, and Hugh Wylam. "Group Tests of Cognitive Development Ideals and a Realization." Journal of Research in Science Teaching, 18(2): 157-168, 1981.
Abstracted by EDMUND A MAREK. 51

Coulter, David, Harvey Williams, and Henry Schulz. "Formal Operational Ability and the Teaching of Science Processes." School Science and Mathematics, 81(2): 131-138, 1981.
Abstracted by DONALD E. RIECHARD. 55

Holliday, William G. "Selective Attentional Effects of Textbook Study Questions on Student Learning in Science." Journal of Research in Science Teaching, 18(4): 283-289, 1981.
Abstracted by EDMUND A. MAREK 59

Boulanger, David F. "Ability and Science Learning: A Quantitative Synthesis." Journal of Research in Science Teaching, 18(2): 113,121, 1981.
Abstracted by DAVID L. DUNLOP 63

Boulanger, F. David and Barbara K. Kremer. "Age and Developmental Levels as Antecedents of Science Learning." Journal of Research in Science Teaching, 18(4): 371-384, 1981.
Abstracted by ROSEMARY K. LUND PADILLA. 68

NOTES FROM THE EDITOR:

Attitude research continues to be an area that attracts many science education researchers. Ten attitude studies are contained in the first section of this issue of ISE. One of the popular areas of attitude research is that of attempts to change the attitudes of elementary school teachers, both pre- and in-service, so they are more positive toward teaching science to elementary school pupils as shown in studies by Piper, Piper and Moore, Moore and Piper, Krockover and Jaus, Piper and Hough, Moore, and Piper. Tamir and two different Israeli colleagues (Sever, Silberstein) investigated pupils' attitudes toward the use of animals in biology teaching. Fortner and Testes investigated students' attitudes about marine education.

The second section of this issue contains analyses of published reports in another popular research area: cognitive development and learning. Linn and others examined correlates of formal reasoning. Shayer and his colleagues reported on the development of seven Science Reasoning Tasks for use in assessing the cognitive development of individuals in groups of 20 or more. Coulter and others examined the situation in which teachers teach algorithms for type-problems and then convince themselves that their students have learned because the students memorize the algorithms rather than understand the concepts. Holliday looked at the effect of textbook study questions on comprehension in high school biology classes. Boulanger reviewed studies, published from 1963-1978, in which a correlation between ability and learning in science was reported. Probably as an adjunct to this review, Boulanger and Kremer examined studies of the 1963-1978 period in which age or grade and developmental level were related to science learning for students in grades 6 through 12.

Patricia E. Blosser
Editor

Victor J. Mayer
Associate Editor

ATTITUDES

Piper, Martha K. "The Investigation of Attitude Changes in Elementary Preservice Teachers in a Competency-Based, Field-Oriented Science Methods Course," Piper, M. K. and K. D. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--*Attitudes; Educational Research; Elementary School Science; *Field Experience Programs; Preservice Education; *Methods Courses; Performance Based Education; Science Education; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Michael J. Padilla, University of Georgia.

Purpose

The purpose of this study was 1) to assess attitude changes in elementary preservice teachers toward science and science teaching that were attributable to a competency-based, field-oriented science methods course; and 2) to assess attitude change of the same students toward science and science teaching one year after completion of the course.

Rationale

Teachers either directly or indirectly determine what is to be taught in elementary classrooms. Teacher attitudes toward science and science teaching therefore might greatly affect the nature of the student's learning experiences. Since it is often a student's previous contacts with science which are cited for later dislike for the subject, it follows that courses which train teachers to teach science should have changing attitudes as a high priority.

Research Design and Procedure

Thirty-six students were randomly selected from a population of 160 preservice elementary teachers in four sections of an elementary science methods coures. The course stressed three major components:

1) Being taught on campus as the students were expected to teach;

2) Participating in field experiences which provide practice teaching entire classes; and

3) Being informed of behaviors expected of them in teaching science to children.

A semantic differential instrument was used to assess attitude changes. Students responded to twelve pairs of bipolar adjectives for each of five protocol phrases. The phrases used in this study were:

1) Science courses in high school;

2) Science as remembered in elementary school;

3) Science education methods courses;

4) Science in the news; and

5) Teaching science to children.

The semantic differential was administered on four occasions: a pretest given prior to the methods course; a second time one month later between the campus and field experience phases of the course; a third time following the entire course and a final time one year later. Using Campbell and Stanley notation the design is as follows:

O X O X O . . . O

The data for each protocol phrase were subdivided into groups called evaluation, potency and activity. Each subset as well as the total semantic differential score was subjected to a one-way analysis of variance comparing pretest scores to the various posttest scores for each of the five phrases.

Findings

When comparing preservice teachers' attitudes on the pretest to those following the campus activities, the author found significant differences for all subsets of items and the total semantic differential score relative to the phrase "science education methods courses" ($p < 0.01$). One subset of items (potency) showed significantly more positive post campus attitudes on the phrase "teaching science to children" ($p < 0.001$) and one subset of items (evaluation) proved significantly more negative on the phrase "science in high school" ($p < 0.02$).

When comparing the preservice teachers' attitudes on the pretest to those following the entire course, the author found significant differences for all subsets of items and the total for the phrase "Science education methods courses." Two subsets of items (potency and activity) as well as the total score were significant for the phrase "Teaching science to children."

No data are reported comparing the pretest scores to the scores obtained one year later, although the author states, "these attitudes are still positive." Only 23 of the original 36 subjects responded at this time.

In a final analysis the author compares male and female responses on the pretest and on the posttest following the entire course. Significant differences favoring females on the pretest were found for all subsets of items and on the total score ($p < 0.03$) for the phrase "Science in high school." One subset of items (evaluation) ($p < 0.02$) and the total score ($p < 0.05$) were significant for the posttest, again showing females to have more positive attitudes.

Interpretations

Through participation in the specially designed methods course, students' attitudes toward teaching science were positively changed and stayed changed over a year's time. There was a difference between the attitudes of males and females "toward science" both before and following the courses. This difference showed males with a more negative attitude. Both sexes "appear to hold attitudes that interfere with science teaching but these attitudes can be changed in spite of past experiences."

ABSTRACTOR'S ANALYSIS

Measuring course effectiveness via attitudinal measures is an important component of an overall course evaluation. This study attempts to document the attitude change of a group of elementary preservice teachers as they progressed through a special science methods course and into their first year teaching. Many questions and some suggestions for improved design and improved accuracy of reporting become apparent.

The most obvious difficulty is the description of the semantic differential and how results of this measure should be interpreted. The author did not describe the subcategories (potency, evaluation, activity) of the instrument. What relation do they have to the protocol phrases? What does a signficant difference in one of the subcategories mean? Should it be interpreted in a special way? Any reader not very familiar with the semantic differential needs this information in order to make independent judgments regarding the author's data.

Also difficult to understand is the choice of the five particular protocol phrases. What do "Science in the news" or "Science in High School" have to do with an elementary science methods course? How do you interpret significant differences for either phrase? Was the semantic differential with these particular protocol phrases a valid way of measuring preservice teachers' attitudes toward teaching science? The author gives no answers for these questions.

Two design questions also arise. It seems reasonable that no test-retest changes due only to familiarity with the instrument were apparent, mainly because of the nature of the instrument itself. However, by administering the semantic differential as a pretest as well as in the middle of the treatment (between campus and field-based course components), it is quite possible that the measure became part of the treatment. Perhaps the instrument sensitized the subjects to think about the importance of "Teaching Science to Children" or to think about "Science in High School." The second problem is the loss of 13 subjects for the final test administration one year after the course ended. How could this mortality have affected the results? How did nonrespondents differ from respondents? No discussion of these important points is made.

The author used several one-way analyses of variance comparing different tests results. In fact, the design involved collecting data at several times throughout the experiment. Thus a repeated measures analysis of variance might have been a more sensitive analysis for finding differences that in fact did exist.

The reporting of the results also gives cause for many questions. The author interprets "trends" as well as significant differences. What is a trend? If an appropriate statistical analysis shows no significant difference within a set of data, then the author is stating that any absolute differences might well be due to chance. When the probability levels are 0.09 and 0.67 for two different protocol phrases, are these both interpreted as trends in the data? The author, in fact, does report both as trends. In another place the data on tables does not reflect the explanation of results in the text. The author states that "attitudes toward Teaching Science to Children were positive in all three subsets." Table 2 shows only two of the three subsets to be significant. Perhaps this is only a proofreading problem but it can be very confusing to a reader. Another issue with the results is the lack of specific data on the attitudes of the subjects one year later. The author states that their attitudes were still positive, but no means or significance levels are reported. Are there still significant differences on certain protocol phrases? Which ones?

One apparent miscalculation arises from the table reporting sex differences on the pre and posttests. The scores for males and those for females on the pretest should, when multiplied by the number of subjects of each sex, added together and subsequently divided by 36, equal the mean score of all subjects on the pretest. For example, 5 males X 44.39 + 31 females X 33.40/36 = 39.93. This does not match the reported mean of 42.86. If, however, the same procedure is used by the reported means for males and females are switched, the scores do match. For example, 5 males X 33.40 + 31 females X 44.39/36 = 42.86. All of the entries can be switched this way and in each case it results in a correct match. Thus it appears that the author has reversed the data for males and females. Conclusions which speak to significantly better attitudes of females over males do not seem accurate. In fact, the opposite appears to be true.

One conclusion appears to be stated in terms that are a bit more general than the data support. The author states that "both men and women appear to hold attitudes that interfere with science teaching...." In fact, no data were collected on this question and the author does not refer to any criteria upon which this statement might be based.

Most studies have the potential for adding to the body of existing knowledge concerning a specific problem. Certainly, this work may be one which has something significant to report. However, several major flaws in reporting leave the reader to wonder about what really occurred during this study. The most severe problems are apparent errors in the data tables, conflicting tables and text and lack of necessary detail in the reported results and procedures. Because of these problems, the reader is left to conclude that no important information was gained through this piece of work. A more well-written report might have circumvented this conclusion.

Tamir, Pinchas and Efrat Sever, "Students' Attitudes Towards the Use of Animals in Biology Teaching," The American Biology Teacher, 42(2):100-103, February, 1980.

Descriptors--*Animal Science; *Biology; Biological Sciences; Educational Research; Science Education; *Secondary Education; Secondary School Science; *Student Activities; *Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Paul Joslin, Drake University.

Purpose

The study had two main purposes:

1. To examine opinions and attitudes of students towards various aspects of using living animals in studying biology; and
2. To identify variables that affect student opinions.

No hypotheses of the cause-effect type were tested. The following observational level hypotheses were examined:

1. Students prefer studying animals to studying plants and microorganisms.
2. Use of animal investigations develops attitudes: some positive, some negative.

Rationale

It is generally assumed that the study of biology for students in grades seven through ten should include the study of plants, animals, and microorganisms. Most text books treat all three, and most programs include observations and investigations with all three. However, there is evidence that students prefer to study animals. Study may mean "read about" or it may mean "observe and investigate." With respect to the latter definition, teachers apparently prefer to use plants. Studies with the Biological Sciences Curriculum Study materials indicate that student motivation drops while studying plants and picks up when returning to the study of animals.

If there are widely-held preferences by students and teachers, and if these affect student motivation, achivement, and attitudes towards experimentation, these should be identified and considered in the design of curriculum and instruction.

Research Design and Procedure

This investigation was based on one questionnaire of three parts given in June, 1978, to 126 students in three schools in Jerusalem, one secular (N = 75), one religious-academic (N = 18), and one occupational (N = 33), with grade distributions as follows: 7-8 (N = 39), 9 (N = 17), 10 (N = 70).

Part A of the questionnaire contained 10 statements to which the subjects responded on a 5-point scale from value 1 = "don't agree", to value 5 = "fully agree." Item 1, "I like to study biology."

Part B contained three multiple-choice items. The first item was a list of animals and students were asked to indicate which they would use in experiments involving irreversible damage. Animals were common and familiar ones from the following groups: Invertebrates, Amphibians and Reptiles, Birds and Mammals.

The second item described a simple investigation in which the compensatory movements of fish are observed after removal of fins. Students were asked their attitudes concerning several operational options.

The third item posed an experiment involving removal of an organ from one mouse and its implantation in another mouse. Students responded to questions eliciting their attitudes towards actually performing the experiment.

Part C asked students to respond to the following open-ended questions: "1. What do you think about the use of living animals in experiments and dissection in research institutions? and 2. What do you think about the use of living animals in the experiments and dissections while learning biology in school?"

Questionnaire results were tabulated and displayed with calculations of averages, percentages, standard deviations and F-ratios.

Analyses included conclusions of statistical significance, and other conclusions of interest, and inferred reasons for both.

Sample size and population distribution were acceptable for the conclusions reported. Techniques of analysis were appropriate and adequate.

Previous research reports were used and, while the investigation was not tightly structured, the conclusions are sound and can form the bases for further research.

Findings

Primary Findings

The investigators reported the following general findings:

1. Most students clearly favor use of live animals as part of their school studies. They view such experiences with animals as increasing both motivation and efficiency in learning and retention.

2. As personal feelings and empathy diminish, logical considerations play an increasing role in decisions regarding experiments with animals.

3. Students have a high level of personal involvement regarding different aspects related to the use of live animals in school.

Secondary Findings

1. Occupational school students are less interested in biology than are students in academic schools.

2. Practically all academic school students are interested in studying biology.

3. Students disagree with the statement, "The study of plants is more interesting than the study of animals."

4. Seventh and eighth grade students find experiments with animals less important than do older students.

5. With respect to taxonomic groups, students:
Agree to use experiments involving irreversible damage to invertebrates.
Prefer not to use animals in the bird and mammal groups and emphatically reject sacrificing pets and farm animals.

Conclusions

The investigators reached the following conclusions:

1. Animals should be used in school but teachers should be aware of, and sensitive to, both positive and negative aspects of experiences with live animals.

2. Use of questionnaires, such as the one used in this study, may serve as a basis for class discussion of important issues related to the use of live animals by students.

Interpretations

From the findings and conclusions reported by the investigators, the following inferences seem reasonable and appropriate:

1. While biology teachers may prefer use of plants to animals in observations and investigations, animals should be used because of the higher motivation to be expected and because positive attitudes may also result.

2. When animals are used, teachers must be aware of, and sensitive to, the attitudes of the students towards the use of animals in investigations especially when irreversible damage is likely.

ABSTRACTOR'S ANALYSIS

The inter-play between student attitudes and motivation, and the effect of these upon achievement, however defined, whether in terms of knowledge, skills, and/or attitudes, is of great interest to teachers and curriculum writers. Of current concern is the high percentage of students who appear to be "turned off" from the study of science and only take required science courses in grades 7-12. Thus, any studies that can reveal and validate possible ways to improve student motivation in science education are desperately needed.

This study focused on students' likes and preferences (defined as attitudes) with respect to the use of animals as objects of study, observation, investigation, and in experiments subjecting animals to harm, and possibly irreversible damage. It revealed that individual students have a high level of personal involvement regarding different aspects related to the use of live animals in school. This finding coupled with the other findings regarding general effects on motivation from the study of plants and animals indicates that studies should be continued in the area.

If the typical high school biology course is the last formal science course taken by most students, then it is crucial that teaching in this course be highly effective. Certainly, the lasting effects, if any, on attitudes towards science generally, and specifically towards biologically based research in agriculture and medicine ought to be sought. This investigation opens up some avenues of research in those directions.

There are also the possible effects of the findings of this investigation on the whole science education program in grades 1-12 but especially in grades 6-12. For example, are the attitudes revealed in this study stable across grade levels? And, whether stable or changing, what inferences may be drawn regarding content and sequence of biological sciences in these grades?

A previous study by Tamir and Jungwirth reported a negative correlation between achievement and attitude in studying biology and zoology. That is, students achieved higher in biology but liked it less and achieved lowest in zoology and liked it more. Clearly this is a problem that ought to interest some science educator just to know what's going on, aside from its possible implications for program development and teaching.

For students of child development and its relationship to science education there is the question of the source or cause of the higher interest in animals. What part, if any, may be due to:

> Keeping of pets and attendant learning about animals?
> Children's stories about animals?
> Lack of children's stories about plants?
> The motions of animals which make them appear more live and evidences for death more striking?

In this same study, Tamir and Jungwirth found that both students and teachers found botany material boring in 7th and 8th grade courses. Why should this be so? What can be done to make botany more interesting to study, given that teachers, for whatever reasons, prefer to use plants for investigations?

Beyond these questions, a major question still remains. Given the need, especially in the junior high grades, to teach essential knowledge, skills, and attitudes, and to do so effectively, what balance, what possible sequences in the use of plants, animals and microorganisms will result in optimal learning? If students prefer studying animals, which particular ones should be used as determined from the needs of teachers, students and the curriculum?

Overall, results of this study as reported appear valid, reliable, and appropriate, make a contribution to the field, and generate questions for further needed research.

Piper, Martha K. and Kenneth D. Moore. "The Effect of a Physics Course for Elementary Teachers on Attitudes Toward Science of Preservice Elementary Teachers," in Piper, M. and K. Moore (Eds.). Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, Ohio State University, 1977.

Descriptors--*Attitudes; College Science; Educational Research; *Field Experience Programs; Higher Education; Physical Science; Preservice Education; Science Education; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Warren J. Safter and James R. Okey, University of Georgia.

Purpose

The purpose of the study by Piper and Moore was to assess science attitudes of perservice teachers who had taken a competency-based, field-oriented methods course. In addition, attitudinal differences were investigated between two subsets of the preservice teachers: those that had taken the course, Physics for the Elementary Teacher, and those who had not. It was hoped that the results of this investigation would have implications concerning science teacher education programs.

Rationale

The researchers recognize the concern for teachers' attitudes toward science as addressed by the Commission on Science Education of the American Association for the Advancement of Science and the improvement of these attitudes. They also consider the effects on attitudes toward science and science courses especially developed for preservice elementary teachers.

Research Design and Procedure

The researchers developed a 26-item Likert-type science attitude scale and mailed it to 160 elementary student teachers graduating in the spring of 1975 from the University of Houston. Of the 99 usable forms that were received, 30 were randomly selected and, with Hoyt's analysis of variance and a factor analysis, used to estimate the reliability and validity of the scale. The responses of the remaining 69 students on each item of the scale were tabulated as well as a total mean for the instrument. Item and grand means were also computed the two subsets of students in the investigation (those students who had and had not taken the physics course). One-way analyses of variance were conducted to test for significant differences between these two group means.

A one-shot case study (Campbell and Stanley, 1966) pre-experimental design was used for the first portion of the study that looked at attitudes of all participating students. A static group comparison was used for the second portion of the study to compare the students who had enrolled in a physics class with those who had not. The results were

analyzed on the basis of mean responses to individual items as well as groups of items with a common focus.

Findings

According to the researchers, the findings are as follows:

1. The competency-based, field-oriented science methods course resulted in a favorable attitude toward science teaching and instilled a desire to teach science through science activities rather than through the utilization of a science textbook.

2. Since the means on items relating to their cooperating teacher's teaching of science fell mid-range, the preservice teachers developed (protective feelings" toward their cooperating teachers.

3. Since the means on items relating to preference among the various methods courses they encountered were mid-range, the preservice teachers "Could not decide which of their methods courses they like best."

4. On comparison of the physics and nonphysics groups, only three items proved significant and the physics group showed more positive attitudes on all three.

5. On comparison of the two groups on items relating to the teaching of science, the physics group "had a more favorable attitude toward activity, hands-on science teaching," and also believed that "elementary children should learn scientific facts in school."

Interpretations

The researchers interpret their findings to imply that attitudes of student teachers toward science can be enhanced by involving them with direct, hands-on science teaching experiences and that methods courses should utilize this approach. They also feel that the evidence indicates preservice teachers believe that "scientific facts are important in teaching science to children" and that science courses specifically designed for preservice teachers can improve their attitude toward science.

ABSTRACTORS' ANALYSIS

There are a number of points about this paper that warrant discussion. Four headings are used below to cluster the discussion points.

Design Considerations. The authors use a one-shot case study and a static group comparison (Campbell and Stanley, 1966) in their study. The situation allows no other choice of design. However, the authors need to exercise appropriate cautions in interpreting data when such designs are used. No data are given about the students who enroll in

the physics class. They respond differently to some of the items on the instrument and we need to decide whether this is because they completed the physics course or for some other reason. If a physics course appeals to certain students it may be that their attitudes are different from their peers before taking the course.

Use of a one-group study with a treatment and posttest (one-shot case study) does not allow cause and effect determinations. The conclusions by the authors that the competency-based, field-centered course "resulted in" certain attitudes is suspect because the favorable attitudes could have been caused by other factors or held by the students prior to their participation.

Instrumentation. A total of 99 usable responses from 160 mailed out contain the data to be analyzed. Thus a substantial portion of the persons who were enrolled in the methods course are not accounted for. Who are they? Were the persons most favorable the ones likely to respond? The authors should provide some evidence to show that their respondents are not a biased sample of all students in the program.

A factor analysis was apparently done with only 30 of the response forms as a check on the validity of the instrument. Most factor analysis texts recommend a much higher number of subjects for analyzing an instrument of this length.

Why are only 69 of the 99 usable returns included in the data analysis? Usually data are compiled from all available responses. In this case the degrees of freedom are reduced for analyzing all the item responses because only a portion of the available responses are included.

The title of the instrument (Piper Attitude Toward Science Scale) and the title of the article both provide the reader with the expectation that attitudes toward science are being examined. Analysis of the 26 instrument items, however, reveals few items (perhaps only one) that deal with attitudes toward science. Instead, the instrument seems to focus on attitudes toward science methods instruction and science teaching.

Several of the 26 items on the scale relate to neither attitudes toward science nor science teaching. Four items deal with feelings toward other methods classes (math, reading, language arts, and social studies) and five others deal with the student teachers' perceptions of the cooperating teacher. Thus 9 of the 26 items on the attitude assessment instrument appear to either not assess attitudes or to assess attitudes other than those related to science or science teaching.

Treatment of Data. Several of the interpretations of the data made by the authors can be questioned. Attributing causal relationships in a one-group study has been dealt with previously. A second possible misinterpretation involves the responses to the five items about how often the cooperating teacher teaches science and the five items about the degree to which the student teachers liked each of the five methods courses they completed. The authors note for each cluster of responses that the item means cluster around 2.5 which is the midpoint on the 1 to

4 scale. From this they conclude that the student teachers are protective of their cooperating teachers and could not decide which of their methods courses they like best.

Another interpretation of the data seems as likely. If one respondent rates a set of four items 4-3-2-1 and a second respondent rates them 1-2-3-4, the means will come out right at the middle of the scale. But neither person could be said to have been unable to make up his or her mind. If dispersion statistics as well as mean scores had been given it would be possible to tell which is the better interpretation.

The authors state that the response to the item--I would need expensive materials to teach science--by the no-physics students is surprising because they feel the need for expensive materials. The no-physics group does score significantly higher than the physics class but their mean response is 1.7 on the 4-point scale. This means that they selected a response (on the average) between "strongly disagree" and "disagree" and therefore doesn't show that they favor use of expensive materials.

A misinterpretation of the response to the item--Elementary children should learn scientific facts in school--is similar to the one just discussed. The authors conclude that the physics group was introduced to more scientific facts than the no-physics group and therefore placed more emphasis on facts in teaching science. There are several problems with this conclusion. First, the two groups did not have significantly different responses to the item. Second, the mean responses were 2.69 and 2.87 for the no-physics and physics group, respectively, on the item which means that they chose, on the average, a response between "disagree" and "agree." Third, it seems unlikely that the no-physics group had no science courses in their training. If physics instruction was thought to influence the one group wouldn't the science courses taken by the other group have influenced them?

Findings and Conclusions. The authors summarize by saying that student teachers' "attitudes toward science can be enhanced" by involving them in hands-on science teaching. The evidence for this is not presented in the paper. The designs used don't allow causal conclusions to be drawn and the attitudes assessed are much more concerned with science teaching than with science.

REFERENCE

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

Moore, K. D. and M. K. Piper. "Factors Underlying Student Teachers' Attitudes Toward Science in a Preservice Elementary Program," Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--*Attitudes; Educational Research; Elementary School Science; Field Experience Programs; *Preservice Education; *Science Activities; Science Education; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Kenneth G. Tobin, University of Georgia (on leave from the Mount Lawley College, Perth, Western Australia).

Purpose

The purpose stated by the authors was to identify factors underlying student teacher attitudes towards activity, hands-on science teaching.

Rationale

The underlying rationale is that science methods and science courses are not the only factors that affect preservice elementary teachers' attitudes toward science.

Research Design and Procedure

Ninety-nine preservice elementary education majors who had participated in a field-oriented science methods course responded to the Piper Attitude Toward Science Scale (PATSS). Responses for each of the 26 items were weighted on a 1 through 4 scale. Strongly Disagree was assigned one point and Strongly Agree was assigned four.

A principal components analysis followed by varimax rotation was used to identify components underlying the item responses.

Findings

Five interpretable factors were identified which accounted for 50 percent of the total PATSS variance. The first factor, which accounted for 20.1 percent of the variance, was concerned with liking and teaching science activities.

A second factor was related to the amount of science taught by the student teacher's cooperating teacher and the cooperating teacher's satisfaction with science teaching. The third factor pertained to the student teacher's past feeling towards science and science teaching. The fourth factor was concerned with activity experiences received in language arts and social studies methods. The final factor was concerned with mathematics activities.

Interpretations

The authors concluded that hands-on science teaching was related to the five factors identified in the analysis. They advocated that researchers direct attention to the provision of improved science experiences during student teaching and improved experiences in language arts, social studies and mathematics. The authors hypothesized that educators might do more to improve attitudes towards hands-on science teaching by directing attention towards areas other than science methods and science content courses for elementary teachers.

ABSTRACTOR'S ANALYSIS

Generalization was limited by the sample used in the study and the validity of PATSS. The authors provide few details on the sample which was claimed to be representative of preservice elementary teachers. Indeed it can be inferred that the sample is the same as that described more fully in an earlier investigation reported in the monographs (Piper and Moore, 1977). If this is the case, the sample may not be typical of preservice elementary teachers because only 61.9 percent of a polled sample returned usable data.

Information was not provided on the validity and reliability of PATSS, and the reader is referred to an earlier investigation reported in the monograph (Piper and Moore, 1977). If this is the case, the sample may not be typical of pre-service elementary teachers because only 61.9 percent of a polled sample returned usable data.

Information was not provided on the validity and reliability of PATSS, and the reader is referred to an earlier investigation reported in the monograph (Piper and Moore, 1977) to obtain necessary data. The high reliability of PATSS (intraclass coefficient = 0.9) provided an indication of the dependability of the measure to differentiate student performance. Although validity was mentioned, no evidence was provided to support the validity of PATSS data.

No rationale was provided for the use of principal components analysis in this investigation. Although a sample of 99 students is satisfactory for a principal compnents analysis, a larger sample is needed if confidence is to be placed in the stability of the solution obtained in this study.

A major problem encountered in reviewing this investigation was that the procedures used in data anlysis were inadequately described and very little data were actually reported. For example, the criteria used to determine the number of factors that comprised the final solution are not given. In an analysis of this type the reader is entitled to know: (1) whether the Kaiser criterion (Kaiser, 1960) or a scree test (Cattrell, 1966) was used to decide on the initial nine-factor solution; (2) the criteria used to determine that five factors represented the best solution; and (3) whether the data were reanalyzed when it was decided to select a five-factor solution in preference to a nine-factor solution.

The greatest problems arose from the lack of data included in the report. The only data table contains factor loadings on seven items on the first factor. To gain an insight into the nature of the five factors, or to judge the adequacy of the conclusions, it is necessary to know the loadings for each item on each factor.

Examination of the 26 PTSS items indicated that 22 were concerned with attitudes to science or science teaching. One item was provided for mathematics, reading, language arts and social studies methods, respectively. Since there was only one mathematics item included on PATSS (I liked math best of my methods courses), it is difficult to understand how a factor could be termed Mathematics Activities. No evidence was provided to support the authors' conclusion that mathematics experiences have a detrimental effect on teaching hands-on science.

At the intuitive level it can be agreed that many factors are likely to affect student teacher attitudes towards activity, hands-on science teaching. The results of this investigation have not contributed a great deal to indentification or clarification of those factors.

REFERENCES

Cattell, R. B. "The Scree Test for the Number of Factors." Multivariate Behavioral Research 1: 245-76, 1966.

Kaiser, H. F. "The Application of Electronic Computers to Factor Analysis." Educational and Psychological Measurement 20: 141-151, 1960.

Piper, M. K. and K. D. Moore. "The Effect of a Physics Course for Elementary Teachings on Attitudes Toward Science of Pre-Service Elementary Teachers." In M. Piper and K. Moore (Eds.), Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Krockover, Gerald H. and Harold H. Jaus. "A Comparative Analysis of Teacher Attitudes in Two National Science Foundation Implementation Programs." In M. Piper and K. Moore (Eds.), Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--*Attitudes; Educational Research; Elementary School Science; *Elementary School Teachers; *Inservice Education; Science Activities; Science Education; Science Instruction; Secondary School Teachers; *Verbal Communication

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Kenneth S. Ricker, University of Georgia.

Purpose

The investigators indicated that they sought to determine "if there were differences in elementary teachers' attitudes toward 1) working with and understanding secondary teachers, 2) teaching at various grade levels, 3) workload and 4) team teaching." The investigation went beyond the stated purpose as indicated by the title of the article. The investigators compared the attitudes of elementary teachers who participated in two different NSF Instructional Improvement Implementation Programs.

Rationale

A lack of research in science education on the communication between teachers was indicated. For this study the researchers assumed that a training program in which elementary and secondary teachers worked together would improve the communication between these two groups, would increase the perception of elementary and secondary teachers' problems, and enhance the development of K-12 science curricula.

Research Design and Procedure

A pre-experimental design that uses a static group comparison was utilized. A summer NSF Implementation Project, having 27 elementary teachers, served as the control group. A NSF Implementation Project conducted in a different summer, having 26 elementary and 20 secondary science teachers, served as the experimental group. In the experimental group an elementary and a secondary teacher were teamed together to teach a group of ten elementary pupils for one week. New teams were formed each week for three weeks. In the control group small groups of elementary pupils were taught by teams composed of only elementary teachers. At the end of each four-week summer program the Science Teacher Opinion Scale (STOS) was administered to the teachers. The STOS, developed by the investigators, is a 15-statement opinionnaire. The content validity was established by a panel of four science educators. A Pearson product-moment coefficient of $r = 0.85$ was obtained, using a test-retest method. Fifty-four elementary and 15 secondary teachers who were not project participants were used to

establish the reliability. Only the STOS scores of the elementary teachers were used in the data analysis. A one-way analysis of variance with two treatment levels was used.

Findings

There were significant differences between the mean scores for the two groups of elementary teachers on six statements. On three statements the teachers in the experimental group disagreed more strongly: 1) Since many elementary teachers are responsible for all subjects, they should be paid more than secondary science teachers, 2) Secondary science teachers have fewer discipline problems than elementary teachers, and 3) Elementary teachers work harder than secondary science teachers. On the other three statements, the experimental teachers disagreed less strongly: 1) A secondary science teacher could teach effectively in an elementary science classroom, 2) I would enjoy team teaching with a scondary science teacher, and 3) Healthy communication between elementary and secondary science teachers is an impossibility. It should be noted that the mean scores for both teacher groups on these six statements were all below 4.0 on a Likert-type scale consisting of five categories. A score of one was strongly disagree and a score of five was strongly agree. The scores for these six statements fell within the categories of undecided, disagree and strongly disagree.

While it was not indicated in the text, the table shows that the mean scores were in the agree categories for only one statement: I would enjoy team teaching with a secondary science teacher. The mean score for both groups of teachers was above 4.0 for this statement.

Interpretations

The investigators suggest a training program which involves the mixing of elementary and secondary teachers can favorably influence the attitudes of elementary teachers toward the working conditions and teaching abilities of secondary teachers. Instructional programs that involve both levels of teachers may be a viable way to encourage interaction and communication between elementary and secondary science teachers. The improvement of intergrade level communication may enhance the development and implementation of articulated K-12 science curricula.

ABSTRACTOR'S ANALYSIS

Communication between intergrade level teachers is important in an educational setting. Effective communication is a two-way process. This study examined one part of the process, that is the attitudes of elementary teachers toward the secondary science teachers. What about the attitudes of secondary science teachers toward elementary teachers and the elementary grades? Would intergrade training programs affect the attitudes of secondary teachers in a way that would also enhance communication? Is there a difference in the opinions of elementary and secondary teachers?

The elementary teachers in this study disagreed with the statement: The open classroom concept is not workable at the secondary school level. Would secondary science teachers also disagree with this statement, regardless if they participated in a training program that mixed elementary and secondary teachers?

Only elementary teachers who participated in NSF projects were used in the study. Are these teachers representative of the total population of elementary teachers? How does one define an elementary teacher? Some elementary teachers teach only science whereas other elementary teachers teach all subjects in a self-contained classroom. Are the opinions of a first grade teacher in a self-contined classroom toward a secondary school science teacher the same as a fifth grade teacher who teaches only science in the elementary schools? One might also ask what are the attitudes of this first grade teacher towards the fifth grade teacher, and vice versa.

If the attitudes of teachers are favorably influenced as a result of a four-week training program, what happens when the teachers return to their respective classrooms? Would the difference in the opinions of the two groups of teachers used in this study exist after one year?

A discussion on the development of the STOS opinionnaire was warranted. How were the 15 statements identified? What criteria were used to select these statements and not others? The fact that only two of the 30 mean scores were in the agree category should have been discussed.

Effecting the process of communication between teachers as a means of helping schools have a K-12 articulated science curriculum represents a large and complicated domain of research. There are a multitude of factors which need to be identified, defined, isolated, and examined in a variety of settings to allow the results to be generalized to the total population. This study represents an initial attempt to examine a factor that may influence the process of communicaton between teachers.

Piper, Martha K. and Linda Hough. "Attitude Changes of Classroom Teachers Cooperating with the Field Component of an Elementary Science Methods Course." In Piper, M. and K. Moore (Eds.), Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--*Attitudes; *Cooperating Teachers; Educational Research; Elementary School Science; Inservice Education; Preservice Education; Science Activities; *Science Education; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Joseph P. Riley, University of Georgia.

Purpose

The authors' stated purpose was to explore changes of teachers' attitudes toward teaching science by having them observe science activities taught in their classrooms by preservice elementary teachers enrolled in a science methods course.

Rationale

Inservice programs have been successful in improving teacher attitudes toward science and science teaching. However, little research has been done to explore other ways of improving these attitudes.

Research Design and Procedure

The attitudes toward science and science teaching of 13 classroom teachers were tested before and after a six-week period. During this time the teachers cooperated with 36 elementary science methods students who taught activity hands-on science lessons in their classrooms as part of a field experience.

Using Campbell and Stanley notation (1971), the design of this study may be represented as One-Group Pretest-Posttest Design.

O X O

Attitudes were measured using a semantic differential. A 12-bipolar set of adjectives with a seven-step scale was used to determine perceptions of five protocol phrases. No information was provided on the validity or reliability of the instrument.

One-way analysis of variance was used to analyze (1) the total responses toward each protocol phrase and (2) the total responses toward each protocol phrase when grouped into the subsets evaluation, potency and activity.

Findings

Significant changes in responses were reported in a negative direction for the protocol phrase, "Science in High School" on the subset activity and in a positive direction toward the phrase "Teaching Science to Children" for the subsets potency and activity.

Interpretations

The authors conclude that classroom teachers' attitudes toward science are more positive following activities taught in their classrooms by preservice teachers. Other comments included under conclusions were:

1) Coordinated efforts between school districts and universities need a greater emphasis in activities rather than inservice workshops to improve science teaching.

2) Preservice teachers need more opportunities to try out science activities in classrooms.

3) Inservice teachers need more opportunities to observe new teaching strategies and activities.

ABSTRACTOR'S ANALYSIS

The problem addressed in this study is one deserving of attention. Evaluation of field-based programs frequently overlook possible spinoff effects such as the programs' influence on the cooperating teachers. However, due to serious weaknesses in this study's design, the research question remains unanswered. The One-Group Pretest-Posttest design employed in this investigation is susceptible to at least five, and possibly six, of the eight threats to internal validity identified by Campbell and Stanley (1971). Chief among these threats are history and testing. History as a validity threat refers to the possible occurrence of many other events in addition to the researcher's "treatment," that offer plausible rival hypotheses explaining the pre to posttest differences. In this study anything science-related, that occurred within the six weeks, could be argued as a factor influencing the teachers' attitudes. Testing itself is a threat to the validity of this study. Pretest sensitization of the teachers may have affected their response on the posttest. This is especially true of attitude measures.

The natural setting of this study placed some obvious limits on the number of experimental designs available to the researchers. Small populations, little or no control of the scheduling of teachers to treatments, and inability to randomize treatment exposures are some examples of the difficulties with conducting research in the schools. The choice often appears to be either the use of an innappropriate design or nothing at all. The application of an inappropriate experimental design because it is the "only one available" does not reduce its inappropriateness. There are other alternatives. A variety of quasi experimental designs are available to researchers who find that

it is not feasible to implement a true experimental design. The use of a quasi-experimental design in this study could have controlled the more serious threats to internal validity and allowed some confidence in accepting the findings.

The report itself does little to increase confidence in the findings or conclusions. The reporting of a significant change in a negative direction toward science in high school for the subset activity is contradicted by the results in the table. In fact, the table indicates just the opposite: a significant change in a positive direction. Other problems concerned with the reporting of the study are: lack of an interpretation of the meaning of the observed changes, reporting of implications as conclusions, and omission of a reference in the bibliography.

The principal limitation inherent in the use of the semantic differential is interpretation. By choosing not to interpret the data, the authors have failed to identify either the theoretical or utilitarian meaning of the differences they observed.

The authors draw a conclusion regarding teachers' attitudes toward science rather than toward the teaching of science which is the independent variable identified in their statement of the problem. The study's other conclusions are unrelated to the investigation. For example, no data were collected on preservice teachers, yet one conclusion is that they need more opportunities to work in classrooms. Findings should be meticulously extracted from the data analysis. Conclusions should then be drawn as to whether these findings support or reject the research hypothesis. Conclusions should not be confused with findings, implications, truths or trivia (Raths, 1967).

The authors have identified a hypothesis worth pursuing. The substantive and stylistic problems addressed in this analysis prevent this investigation from acting as other than a pilot study. This study suggests, rather than supports, the hypothesis that preservice teachers influence their cooperating teachers' attitudes.

REFERENCES

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1971.

Raths, J. "The Inductive Process: Implications for Research Reporting." Educational Leadership 24(4): 357-363, 1967.

Silberstein, Moshe and Pinchas Tamir. "Factors Which Affect Students' Attitudes Towards the Use of Living Animals in Learning Biology." Science Education, 65(2): 119-130 (1981).

Descriptors--Animal behavior; *Biology; Elementary-Secondary Education; *Elementary School Science; *Laboratory Animals; Questionnaires; Science Education; *Secondary School Science; Sex Differences; *Student Attitudes; Zoology

Expanded abstract and analysis prepared especially for I.S.E. by Constance M. Perry, University of Maine at Orono.

Purpose

The investigators' purpose was to identify factors which affect students' attitudes toward the use of animals in research and in learning biology. The investigators did not state specific hypothesis or research questions but listed the following as the specifics they studied.

(1) examined the attitudes of students toward the use of animals in research as compared with their use in school;
(2) studied the extent to which cognitive and affective considerations as well as value judgments underlie students' attitudes toward the use of living animals in learning biology;
(3) examined the attitudes of students to different kinds of animals as objects for experimentation;
(4) identified factors which influence attitudes of students to the use of different kinds of animals in experiments;
(5) examined the factors which affect students' decisions as related to the performance of particular experiments which cause some damage to the experimental animals;
(6) examined differences in attitudes by grade level;
(7) examined attitude differences between boys and girls.

Rationale

The research was based on the underlying assumptions that direct observations and experiments with living organisms are congruent with the general effort to use primary sources in learning a discipline (Blum and Silberstein, 1979; Orlans, 1968; Orlans, 1972) and that there is no clear cut answer to the question: What are the effects of using living animals in the classroom with respect to students' attitudes toward life, living animals, and biology? The study was an outgrowth from two previous studies (Tamir and Sever, 1980; Tamir and Hamo, 1980) in which it was found that most students were interested in studying live animals, most students believed direct observations and experiments were superior to learning from secondary sources, and most students showed concern for and affection toward living organisms, especially higher animals.

Research Design and Procedures

Sample: The sample consisted of 577 students from four grade levels as follows: 5th grade (one class, N=25) 7th grade (nine classes, N=269) 9th grade (four classes, N=109), and 11th grade (six classes,

N=174). Seven schools, in the central part of Israel, were involved and the school selection was done on the basis of convenience. Except for the fifth grade, which was made up of talented students, the classes were heterogeneous with regard to ability. The information about sex of students was available for only 348 students, 180 boys and 168 girls.

Instruments: All students involved in the study completed a four-part questionnaire.

Part A consisted of 30 statements to which the subjects responded on a 5-point scale in which 1 = don't agree and 5 = fully agree. The first six statements were from Tamir and Sever (1980), and another 14 statements were identical to those used by Tamir and Hamo (1980). The remaining 10 items were designed specifically for this study.

Part B consisted of a case study where Dan is described as one who likes animals but in order to study the swimming behavior of fish wishes to cut off the fins. He hesitated to actually cut off the fins, since he realized that he was going to cut organs of a living creature. The student is asked, "what would you do if you were faced with this dilemma? Would you cut off the fins? Give your reasons."

Part C consisted of a list of 20 animals, 10 "lower" (worms, scorpions, flies, ants, honeybee, fish, frog, lizard, poisonous snake, non-poisonous snake) and 10 "higher" (chicken, pigeon, mouse, bat, rabbit, pig, goat, cat, dog, monkey). The subjects had to indicate for each animal if they would be in favor of using it in experiments which cause irreversible damage.

Part D required the subjects to choose one or more out of ten possible responses to the statement: "You are assigned to dissect a mouse and implant in it an organ taken from another mouse."

Data Analysis

The mean scores of Part A were analyzed by two way analyses of variance by grade level and sex, and t-tests were computed between boys and girls by grade level and between subtests. (There were eight subtests, each representing a distinct aspect of use of animals). Frequency distributions were constructed for subtest responses. Also, Duncan's multiple range test was computed for grade levels.

The list of animals in Part C was categorized according to taxonomical level, benefit to man, living site, and general affection of people toward the animal. A stepwise multiple regression was performed. The generated scores were combined into one data file providing a basis for the regression analysis. The five scores, Harmful, Higher, Taxonomic Level, Useful liked and Domestic were used as independent variables. Twenty separate regressions were run with each animal serving once as the dependent variable. The twenty items of Part C were also intercorrelated and submitted to R factor analysis with varimax rotation.

Analysis of variance was computed between grade levels and between sexes for Parts B and D. The reasons given in Parts B and D were classified into five categories and a frequency distribution constructed.

Findings

The following represent the major findings.

Part A

1. Statistically significant main effects were found for grade level ($p < 0.01$) and for sex ($p < 0.01$) as well as an interaction effect ($p < 0.05$). The mean scores indicated that boys' attitudes toward the use of animals become more positive with age, and in general boys have more positive attitudes than do girls. Girls tend to have less positive attitudes in grades 5 and 9 than in grades 7 and 11.

2. There is a significant difference between the results of subtests 2 and 3, (p 0.0001) showing that students differ in their attitudes toward the use of animals in research as opposed to in schools: 79 percent support the use of animals in research, while 64 percent support their use in schools. When animal suffering is involved, 38 percent agree to use in schools, and 62 percent agree for research. In experiments causing irreversible damage, 56 percent support animal use in research and 44 percent in schools.

3. Seventy-four percent believe one learns more from experiments than from books (subtest 5).

4. About two-thirds of the students are in favor of the teacher doing the experiments to reduce animal suffering (subtest 7).

5. Eighty-five percent agree to the use of animals in schools provided the animals are returned to their natural habitat (subtest 7).

6. Sixty-eight percent state any damage to the animal should be temporary and reversible (subtest 7).

7. Sixty-eight percent oppose the view that one can treat animals as he likes (subtest 1).

8. Twenty-nine percent agree to unrestricted use of animals by man (subtest 1).

9. Seventy-one percent agree to the use of animals for the benefit of man if cruelty is prevented (subtest 6).

10. There is a significant difference between the means of subtests one and six (p 0.0001).

11. Duncan multiple range test revealed statistically significant differences between 5th grade students and others.

Part C

1. On the average, in the 20 separate regressions, the five factors chosen explain 56 percent of the variance in students' responses with a range between 32 percent (for frog) and 81 percent (for bat).

2. The results of the multiple regression analysis are ranked from 5 (variable entered first) to 1. Only entries making statistically significant contributions to changes in variance ($p < 0.05$) were included. An importance index for the five factors was computed.

$$\text{Importance index} = \frac{\text{Average Rank x percent of items contributing}}{100}$$

The importance index for the factors was as follows:

Harmful to man = 3.7	Higher animals = 3.3
Useful to man = 2.4	Liked by people = 2.2
	Domestic animals = 1.9

3. Four factors accounting for 61 percent of the total variance resulted from the R factor analysis with varimax rotation.

Factor 1 = "Higher - Useful" (all mammals but bat & mouse, all birds)
Factor 2 = "Lower" (worms, insects)
Factor 3 = "Repulsive" (bat, lizard, poisonous and nonpoisonous snakes, scorpion)
Factor 4 = "Experimental" (mouse, fish, frog)

4. Analysis of responses to the four subtests generated by the factor analysis by grade level and by gender yielded.

   a. There was no significant difference between boys and girls in grade 5.

   b. In grade 7, boys supported the use of lower animals ($p < 0.01$) more than did girls.

   c. In grade 9, 23 percent boys supported the use of higher animals, contrasted with 7 percent of the girls.

   d. In grade 11, significant differences by sex existed beyond the 0.01 level for all subtests.

      Lower: boys 78 percent, girls 53 percent
      Experimental: boys 70 percent, girls 44 percent
      Repulsive: boys 54 percent, girls 29 percent
      Higher-Useful: boys 24 percent, girls 7 percent

   e. 48 percent of 7th graders were ready to sacrifice "repulsive" animals, whereas only 30 percent of the 9th graders were ready to do so.

Parts B & D

1. Twenty-five percent were willing to cut the fins off a fish. (Part B)

2. There were significant differences between grade levels ($p < 0.01$) on the fish dilemma. The 5th graders were the least willing to remove fins and the 9th graders the most. (Part b)

3. There was no significant difference between boys and girls on the fish dilemma. (Part B)

4. Sixty percent were ready to implant an organ in the mouse. (Part D)

5. Girls were on the average significantly less inclined to operate on the mouse than were boys. (Part D)

6. There were no significant differences among grade levels in the mouse dilemma. (Part D)

7. In order to examine the kinds of reasons students offered when given the dilemmas, reasons were classified into categories.

    1. The importance for learning, research and human welfare;
    2. Consideration for animals and their well-being;
    3. Conventions;
    4. Personal attitude related to fear, utility or disgust;
    5. Evading the dilemma by taking resort in alternatives such as film, TV, books, etc.

        a. Of 283 reasons given in the open ended fish fin dilemma 30 percent were category 1, 43 percent category 2 and 27 percent category 5. (Part B)
        b. In the mouse dilemma of 296 reasons, 46 percent were category 1, 27 percent category 2, 15 percent category 3 and 12 percent category 4. (Part D)

Interpretations

Part A

The authors concluded:

"...that student support for use of animals is restricted by moral and emotional considerations."

The attitudes of students represent an intermediate position between the homocentric view which implies that man may use animals with no restrictions and an egalitarian view that all creatures are equal and have similar rights.

"...there is a general trend revealing more positive attitudes toward the use of animals and avoiding cruelty by girls."

"Eleventh grade students were most positive regarding the use of animals in instruction and least concerned to put restrictions on that use."

Part C

The authors concluded:

"Students are much more inclined to use in experiments lower, harmful, wild and hated animals."

"...the most important factor...in deciding about the use of particular animals in experiments is the extent to which this particular animal is harmful to man. The least important factor is whether the animal is wild or domestic."

"...that up to grade 9 there are only small differences between boys and girls in their readiness to use animals in experiments which cause irreversible damage." In grade 11 marked differences exist, with girls considerably more reluctant to sacrifice animals.

Grade 5 students were considerably more reluctant to use animals than the higher grades.

"...the degree of usefulness of damage to man associated with the animal, and the taxonomical position of the animal..." are the two major factors affecting students in their differential attitude toward the use of animals.

Parts B and D

The authors concluded:

"...the older students, value more the learning experience, offer more justifications for using animals and are less concerned about the animal suffering. At the same time they are inclined to prefer alternatives to direct experimentation."

Concern for the well being of animals was much greater in relation to the first, seen as in an instructional context, as opposed to the mouse, viewed as in a research context. "Students make a clear distinction between the use of animals in research as opposed to the use of animals in instruction."

Overall conclusions

Students are ambivalent toward the use of animals for instruction. They favor the use with certain provisions and restrictions. "The implications is that animals should be used whenever their use brings about improved learning, provided that undue suffering and cruelty are avoided..."

"...the attitude of children towards the use of animals in experiments is crystalized between the ages of 12 and 17."

ABSTRACTOR'S ANALYSIS

Students' attitudes toward the use of living organisms in research and in learning biology is an area of research that has received moderate attention. This study adds another dimension to the literature by looking at the factors which affect student attitudes toward the use of animals. By identifying the factors behind the attitudes, better decisions can be made concerning the use

of animals in schools. In addition, by making teachers aware of the underlying factors, they in turn can better prepare their students to use animals. Curricular and teacher planning which takes identified factors into account can promote learning and help in the development of positive attitudes towards biology and animals as they relate to man. Few studies have attempted to look beyond the attitudes to find the factors that affect them. The purpose of the study is indeed a noble one.

The study was a complex one. There were so many parts and subparts that it was difficult to follow what was done or to discern the results. Several alterations could have made it more easily understandable. Limiting the parts of the study reported in one article would have enabled the authors to more clearly describe the treatment and results. Specific hypothesis or research questions would have clarified the scope of the study. Describing the data analysis concisely rather than spreading the description through several pages of results would have enabled the reader to quickly find out how the data were analyzed. The results were not reported in a consistent manner. Sometimes percentages were given without mentioning whether the differences were significant, and when significant differences were found-the level was not always reported. The amount of information in the article seemed to be just too much to deal with in a clear and consistent fashion.

A sample size of 577 is to be commended. Although the schools were chosen on the basis of convenience, because of the heterogenity there is no great reason to doubt that they are representative of students in Israel's schools with the exception of the fifth grade group. The authors state that the 25 fifth graders are all talented students. The reason given for limiting the number of fifth graders is the authors' doubts as to the ability of such young students to respond meaningfully to the questionnaire. Any results concerning the fifth graders, because of the lack of randomness and the small N, should be regarded with caution.

Although the sample size was 577, information about the sex of students was only available for 348 (180 boys and 168 girls). Again, the subsample size is adequate but the question as to whether the distribution of the 348 is fairly equal for all grade levels is not addressed. An unequal distribution by grade level of either or both sexes would cause one to question the significance of findings concerning differences by sex for grade levels. Such information is necessary.

In regard to data analysis, clarification is needed for the analysis of variance procedures. One does not know if the analysis of variance was computed for 2 sexes by 4 grades or 2 sexes by 2 grades.

In Part A when the 30 statement questionnaire is divided into eight subtests, internal consistency for each subtest was stated as relatively high (Cronbach levels between 0.25 and 0.52, while the mean point biserial correlations ranged between 0.47 and 0.78). "Relatively high" is a generous statement. It is recommended that reliability coefficients be at least .70 (Mitzel, 1982). In making conclusions from the Part A findings, the internal consistency values as well as the limited number of items per subtest should be kept in mind.

Several factors were identified as affecting students' attitudes toward use of animals. One factor the authors identified, research as compared to instructional use, may have been magnified. The authors viewed the fish dilemma (Part B) as an instructional one and the mouse dilemma (Part D) as research. Students were more willing to operate on the mouse than the fish, therefore it was concluded that they were more positive toward the use of animals in research than for instruction. The premise that the mouse dilemma was viewed as research by students can be questioned since the authors state that the mouse is one of three animals widely used in school experiments. Also, the mouse dilemma statement is worded; "You are assigned to dissect a mouse and implant in it an organ taken from another mouse." The word "you" would be taken by students to mean themselves and therefore could easily be viewed as in an instructional context rather than research. Perhaps other reasons need to be studied as to why the students were more willing to operate on the mouse.

This study is a beginning in identifying factors affecting student attitudes toward the use of animals and in identifying attitude trends. The findings could be valuable to people working directly with school children; but as was alluded to earlier, the findings need to be more clearly and concisely reported if they are to reach the audience that could benefit from the information. Further research would also be valuable in order to clarify the research versus instructional use issue, the general trend of boys being more willing to use animals than were girls and becoming more willing with age, and the confusing results of girls' willingness to use animals.

REFERENCES

Blum, A., and M. Silberstein. "Trends in Biological Education in Israel." Journal of Biology Education, 13:41-51, 1979.

Mitzel, Harold E. (Ed.). Encyclopedia of Educational Research (5th ed.) New York: Macmillan Publishing Company, Inc., 1982, p. 1600.

Orlans, B. "The Boundaries of Use of Animals in High School Biology." The Science Teacher, 35:44-47, 1968.

Orlans, B. "Five Organisms in High School Biology." The American Biology Teacher, 34:343-345, 1972.

Tamir, P., and A. Hamo. "Attitudes of Secondary School Students in Israel Towards the Use of Living Organisms in the Study of Biology." Int. J. Study Animals Problems, 1:299-311, 1980.

Tamir, P. and A. Sever. "Students' Attitudes Toward the Use of Animals in Biology Teaching." The American Biology Teacher, 42:100-103, 1980.

Moore, Kenneth D. "An Investigation of the Science Background of Elementary Teachers," Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--Educational Research; Elementary School Science; *Elementary School Teachers; *Inservice Teaching; Inservice Teacher Education; Research Methodology; Science Education; *Surveys

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Joseph P. Riley, University of Georgia.

Purpose

The study was designed to assess the status of science teaching and elementary science teachers preparation in science content, methodology, and recency of training.

Rationale

Teachers in the Far West and Great Lakes Regions had been surveyed and found to be lacking in science knowledge and methods.

Research Design and Procedure

A questionnaire was distributed to 260 elementary teachers who were selected at random from an available population of 6,185. This population consisted of teachers employed in 21 school districts in Harris County, Texas. Of the 260 questionnaires, 127 were returned for a response rate of 48.9 percent.

The teachers were asked to supply information regarding (1) the amount of science they taught, (2) the year of their last science methodology course, (3) the year of their last science content course, and (4) the content area of their last science content course.

Findings

The study found that elementary teachers lack preparation in science content and science methodology.

ABSTRACTOR'S ANALYSIS

This study provides the opportunity to observe the effects of violating certain specific demands of descriptive survey methodology.

What happens when a questionnaire is designed that fails to address a specific research objective? The questionnaire becomes an exercise in data gathering. Although the data may allow casual observations and

inferences, it lacks the objectivity, specificity, relevance and suitability for purposeful observations and meaningul interpretation.

What happens when a survey questionnaire fails to ask precise questions? One result may be that a four-item questionnaire, such as the one used in this study, is returned with as many as 53 percent of the respondents failing to respond to all four items. The questionnaire language should be unmistakably clear and precise. Asking teachers to respond to a question using such relative terms as "none, some and a great deal" misses the precision needed for obtaining valid and reliable data.

What happens when a survey fails to be piloted on a small sample? Difficulties with question ambiguity are not detected and may result in a high rate of unanswered questions.

What happens when 15 percent of the respondents fail to respond to a question? Bias enters into the data. Assumptions become extremely tenuous. When 15 percent of the teachers failed to respond to a question asking when they had their last content course, the author offers the assumption that they had no content courses. Another, just as valid, assumption is that they all had courses between 1970 and 1974, increasing this category to 40 percent and indicating a much more optimistic picture than that reported.

What happens when only 50 percent of the surveyed population returns the questionnaire? The researcher is not in a good position to tabluate data and announce results because he doesn't know which half of the teachers has responded. The researcher assumed he has a random 50 percent but provides no justification for why the reader should assume this. Follow-ups are an essential part of a survey plan in order to reduce nonrespondents to as small a percentage as possible. According to Tuckman (1978), "Most researchers are unwilling to accept a return of less than 75 to 90 percent (and rightly so)."

What happens when no follow-ups are sent to nonrespondents? Without follow-ups one can expect between one-third to two-thirds of the questionnaires returned. This can be extended by 10 to 25 percent with a follow-up (Tuckman, 1978).

What happens when a simple random sampling is applied to a multicultural population? Some worthwhile information is lost that might have been picked up by either a simple stratified or proportional stratified sampling techniques. Simple random sampling is most appropriate when the populaton is generally a homogeneous collection of individuals.

The limited number and type of questions used in this survey suggest they were intended for demographic rather than planned research purposes. This might explain their limitations when evaluated in terms of acceptable descriptive survey methodology.

REFERENCE

Tuckman, Bruce W. Conducting Educational Research. Atlanta, GA: Harcourt-Brace-Jovanovich, Inc., 1978.

Piper, Martha K. "Comparison of Attitude Changes of Preserivce Elementary Teachers Toward Science Between a Traditional Science Methods Course and a Non-Traditional Science Methods Course." In Piper, M. and K. Moore (Eds.), Attitudes Toward Science; Investigations. Columbus, OH: SMEAC Information Reference Center, 1977.

Descriptors--*Attitudes; Educational Research; Elementary School Science; *Field Experience Programs; *Methods Courses; Preservice Education; *Performance Based Education; Science Education; Science Instruction

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Leon W. Benefield, Jr. and James R. Okey, University of Georgia.

Rationale

The responsibility of teacher educators to train competent teachers has often not been met because of a lack of field experiences in training programs. The new child-centered science programs require field experience to implement--experience offered more often in the newer "nontraditional" education methods courses.

Chapman (1971) and Strawitz (1976) both report more favorable attitudes of preservice teachers enrolled in activity-centered methods courses.

Research Design and Procedure

Approximately 150 students enrolled in elementary science methods courses at two universities were involved. The students received one of two treatments:

Traditional Science Methods--The courses were lecture-oriented with no field experience. Practice teaching was confined to peer teaching on campus.

Nontraditional Science Methods--The courses were competency-based and field-oriented with 40 percent of the courses being field experiences.

Data were collected with a semantic differential instrument (Osgood, Suci, and Tannenbaum, 1957) using a seven-step scale with 12 sets of bipolar adjectives. The adjectives were grouped into subsets to assess the evaluation, potency, and activity dimensions of attitudes. The subjects responded to each of five protocol phrases:

1. Science in High School
2. Science as Remembered in Elementary School
3. Science Education Methods Course
4. Science in the News
5. Teaching Science to Children

The instrument was administered four times using two different research designs (Campbell and Stanley, 1966).

Administration Time	University	Design	
1	A	O X O	One group pretest posttest design
2	A	X O	Static-group comparison

	A	X O	
3	A	X O	Static-group comparison

	B	X O	
4	A	X O	Static-group comparison

	B	X O	

X - Traditional Science Methods Course

X - Nontraditional Science Methods Course

O - Administration of Semantic Differential

The total and subtest scores (evaluation, potency, and activity) were subjected to ANOVA for each of the five protocol phrases at each administration time to determine whether there were attitude changes or differences.

Findings

Significant changes of differences were found on three of the four administrations of the attitude assessment instrument.

Administration 1--Showed that preservice elementary teachers have a more negative attitudes toward the "Science Education Methods Course" after experiencing a traditional science methods course.

Administration 2--Showed that preservice elementary teachers enrolled in a nontraditional science methods course have a more positive attitude toward the "Science Education Methods Course" and "Teaching Science to Children" than preservice teachers in a traditional methods course.

Administration 3--Showed the same results as Administration 2. In this case, however, the students being compared were enrolled in different universities.

Administration 4--Showed that there were no differences in attitudes toward any of the five protocol topics between students enrolled in nontraditional science methods courses at two universities.

Interpretations

According to the author, preservice elemtenary teachers participating in competency-based and field-oriented science methods courses developed more positive attitudes toward science than did teachers in a lecture-oriented course with no field experiences. Science educators need to devise science methods courses that include stated competencies and field experiences as one way to improve attitudes toward science.

ABSTRACTOR'S ANALYSIS

The researcher used two pre-experimental designs (Campbell and Stanley, 1966) to study questions relating to attitudes of preservice science teachers. Given the difficulty of randomly assigning students to classes, the choice is not surprising. However, these designs do have limitations and threats to the validity of findings need to be addressed. The specific threats to be considered are:

Selection Bias--To what degree are the two groups comparable that participated in the three static group comparison studies? In one of these cases the students are from the same university. However, we are told nothing about their characteristics or how students are assigned to classes so that a judgment can be made the the groups are similar. In the other two comparison studies, students are from different universities and the prospect of having noncomparable groups is even greater. If data on age, sex, SAT scores or other such demographic information were provided, readers would be in a better position to judge whether selection bias is a factor worth considering seriously.

Testing--Assessing changes with studies of only one group offers data interpretation problems. There is a possibility that the original administration of the attitude measure in the one group pretest-posttest design may have sensitized the students to the type of changes being measured. After the test was administered the subjects could have discussed the various items and individuals could have developed an idea of what would be the desired responses. When the measure was given the second time to the same group, these preconceived notions may have affected the reponses to the instrument. If a control group had been available this threat could have been eliminated. History (what happens to the students besides the treatment) may have also been a cause of changes in attitudes.

Aside from design consideration there are a number of points about the study that warrant comment.

1. Because the data from Administration 4 were not available, it is impossible to determine the total number of subjects involved and the number of identifiably different groups in this portion of the study. An apparent typographical error in the heading of Table 3 also makes group membership difficult to discern.

2. The statement that a number of attitude "changes" were identified with Administration 2 and 3 of the measure is not warranted. In the experimental design used by the author only Administration 1 offered information regarding "changes" or "development" of attitudes. It would have been more accurate to report attitude differences between the various groups in the study.

3. In the conclusion is found the statement that "...competency-based and field-oriented science methods courses develop more positive attitudes toward science..." the evidence, however, shows differences on overall scores associated with only two of the protocol phrases--Science Education Methods Course and Teaching Science to Children. Thus, the attitude differences seem to be associated with phrases having to do with learning to teach and teaching science but not science itself. The paper would have been clearer if the areas of attitude difference or change had been defined more accurately.

Studies such as this can offer valuable information to individuals developing or instructing science methods courses. It would be helpful to find out what kind of courses for preservice teachers influence attitudes and which attitudes are influenced. Studies can then be planned to follow these preservice teachers into the classroom and assess the effect of these positive attitudes on teaching.

REFERENCES

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1966.

Chapman, H. L. "The Effect of an Inservice Summer Elementary School Science Education Institute on Attitudes Toward Science and Science Understandings of Elementary School Teachers." Doctoral dissertation, University of Connecticut. Dissertation Abstracts International, 32: 2503A, 1971.

Osgood, C. E.; G. J. Suci; and P. F. Tannenbaum. The Measurement of Meaning. Urbana, IL: 1957.

Strawitz, B. M. "The Effects of an Activity-Centered Elementary Science Methods Course on the Attitudes of Preserivce Teachers." Paper presented at the Forty-Ninth Annual Meeting of the National Association for Research in Science Teaching, April 1976.

Fortner, R. W. and T. G. Teates. "Baseline Studies for Marine Education: Experiences Related to Marine Knowledge and Attitudes." Journal of Environmental Education 11(4):11-19, 1980.

Descriptors--*Achievement; *Educational Assessment; Environmental Education; *Educational Research; Grade 10; *Interdisciplinary Approach; Marine Biology; *Oceanology; Science Education; Secondary Education; *Secondary School Science; Student Attitudes; Surveys

Expanded abstract and analysis prepared especially for I.S.E. by Richard M. Schlenker, Maine Maritime Academy.

Purpose

The stated purpose of this investigation was to measure the level of contact students have had with oceanic activities and information. While the authors did not formulate specific hypotheses regarding the outcomes of their investigation, they did seek answers to the following research questions:

1. How much do the students know about the oceanic environment and its influence on marine affairs?
2. What are the students' attitudes toward some marine issues?
3. Is residential proximity to the coast related to students' levels of marine awareness?
4. What marine-related experiences do the students perceive to be most influential in providing them with oceanic information?
5. What experiences can be shown to be related to marine attitudes or knowledge?

Rationale

The study was based upon the premise that if a given experience or set of experiences were influential in providing information to students having positive attitudes toward marine issues and/or high marine knowledge scores (as evaluated by test instrument), those experiences which were manipulable might be used as means of increasing marine awareness. In addition, experiences consistently correlating with low marine knowledge scores or negative attitudes signaled a possible need to concentrate instruction in such a manner as to overcome the effects of those experiences. The findings of the study, therefore, would provide a decision support system, independent of intuition and guesswork, for those charged with making decisions regarding the implementation of marine education in the school curricula.

Research Design and Procedures

Sample: Seven hundred eighty seven tenth-grade students from thirty Virginia schools served as subjects. The schools were selected for inclusion using a stratified sampling program based upon their proximities to the coast. Fifteen schools were randomly selected from Virginia's coastal area (defined by Virginia's Code as "Tidewater Virginia"). Fifteen randomly selected schools from Virginia's interior were also used. Classes of approximately 30 students each, whose students reflected the entire range of abilities, were administered the test instrument.

Instrument: A modified form of the Survey of Oceanic Attitudes and Knowledge, a three-part instrument, was used to gather data; one section evaluated students' attitudes [4-category Likert scale - (4) Definitely agree; (3) Tend to agree; (2) Tend to disagree; (1) Definitely disagree.] toward the marine environment; another, knowledge about the marine environment; the third evaluated students' experience levels with the marine environment.
Administration: Administration of the evaluation instrument (forms A, B, and C) was completed in one class period of forty-five minutes duration.
Data Manipulations: Several mathematical techniques were used in evaluating different forms of the instrument and data collected through administrations of the instrument:

1. Chi-square analyses were used to check the equivalence between knowledge sections of different forms of the instrument.
2. The differences between mean knowledge scores were evaluated using t-test procedures.
3. Descriptive statistics were used to describe students' understanding of the ocean as a: chemical medium; physical system; biological community; political interface; cultural influence, and threatened resource.
4. Attitude scores were produced by calculating mean attitude item response scores. They were then compared to demographic variables using analysis of variance techniques.
5. A mean total sample attitude scores was produced; attitude means were then compared (tidewater and inland).
6. Attitude response categories were evaluated using percentages.
7. Marine knowledge and attitudes were compared using a Pearson product moment correlation.
8. ANOVA was used to evaluate the effect of demographic variables (race, sex, residence) upon knowledge.
9. Respondents rank ordered ten experience categories according to perceived importance in providing information. Mean category rank scores were used to develop a "Scale of Perceived Influence" of experience upon level marine knowledge.
10. Relationships between marine knowledge and experience categories were examined using multiple regression analysis.
11. The relationship between attitudes and experience categories were examined using multiple regression analysis.
12. Correlations were calculated between specific experience items and knowledge categories and between specific experience items and attitudes.
13. Tukey's multiple range test was used to evaluated the relationship between group means and level of knowledge.

Findings

The following major findings relate item-for-item to the data manipulations cited above.

1. Subject response patterns for items which differed from one test form (A, B, C) to another were the same regardless of the form used.
2. The mean knowledge score (12.9) for the tidewater group differed significantly from that (11.9) of the inland group.
3. The average percentages (from forms A, B, C) of correct answers for instrument items testing subjects understanding of the oceans were;

chemical medium, 53.8; biological community, 46.2; physical system, 47.1; political interface, 53.9; cultural influence, 43.5; threatened resource, 51.0.

4. Based upon a criterion of $P \leq .1$ demographic factors appeared not to affect subjects' attitudes about the marine environment. The lowest probability of any other than chance relationship ($P \leq .271$) occurred for race; white subjects exhibited more positive attitudes.
5. The mean attitude score for the entire sample was 3.06; the possible score was 4.0. Scores between the two groups did not differ significantly.
6. With regard to total attitude scores, 0.89 percent were strongly negative, 6.5 percent slightly negative, 74.8 percent slightly positive, and 17.8 percent strongly positive.
7. The Pearson correlation between attitudes and knowledge about the marine environment was $r = .43$, significant at $P \leq .01$.
8. The effects of demographic factors upon marine knowledge are shown in Table 1.

TABLE 1

Effect of Demographic Factors on Marine Knowledge

Source of Variation	Significance
Main Effects	.001
Residence	.001
Race	.001
Sex	.001
2-way Interactions	.05
Residence/Race	$P > .05$
Residence/Sex	.01
Race/Sex	$P > .05$
3-way Interactions	$P > .05$
Residence/Race/Sex	$P > .05$

9. The results of perceived influence upon marine attitudes and knowledge scores are shown in Table 2.

TABLE 2
Rank Ordering, Perceived Influence of Experiences
Knowledge and Attitudes

Numerical Value*	Coastal Subjects	Numerical Value*	Inland Subjects
7.17	T.V. Specials	7.07	T.V. Specials
5.75	Movies	6.04	Movies
5.62	Magazines	5.52	Ocean Study Class
5.53	Regular School Class	5.49	Magazines
5.47	Beach Recreation	5.43	Regular School Class
5.45	Public Aquaria	5.20	Beach Recreation
5.16	Books	5.12	Public Aquaria
5.15	Education on Beach	5.02	Books
4.93	Newspapers and T.V. News	4.86	Newspapers and T.V. News

* Score range 1-10

10. When marine knowledge was evaluated against marine experience categories (Table 2), 6.6 percent of the variance was related to some common factor.
11. The common variance between attitudes and marine experience categories was similar to the 6.6 percent figure for knowledge and experience.
12. In the comparison, three items individually accounted for 10 percent of the variance; (a) number of Cousteau specials seen on T.V., r = .37; (b) frequency of reading National Geographic, r = .31; (c) swimming ability, r = 31.
13. The range test indicated the greater the participation watching Cousteau specials, reading National Geographic, and swimming, the higher the knowledge scores.

Interpretations

The authors included:

1. Virginia's tenth grade students have a fairly low level of knowledge about the ocean.
2. Students perceive marine-related television programs as the greatest influence on their knowledge of the ocean.
3. Access to marine experiences should be equalized across socioeconomic barriers by providing such experiences in classrooms.
4. An attempt should be made to increase student awareness of the potential of females and non-whites for involvement in marine activities.

ABSTRACTOR'S ANALYSIS

Fortner and Teates (1980) are to be commended for their success in expanding the existing data base reflective of levels of marine awareness possessed by U. S. school children. The major goal of the National Marine Education Association, since its incorporation in 1976, has been to educate, from the

thoughts of Goodwin (1977) and Goodwin and Schaadt (1978), a marine-literate society. Baseline studies of the type discussed here are invaluable to school administrators and teachers alike tasked with the implementation of marine education programs. Education for marine literacy cannot begin until educational points of departure have been established. These data help provide such a point.

The data gathering instrument requires comment. The instrument used was the Survey of Oceanic Attitudes and Knowledge. While the instrument is adequately described, a source from which the instrument may be obtained seems to have been omitted. In addition, the reliability of the instrument must be questioned since one or more such coefficients have not been included.

Suggestions for Further Research

There is a great need on the part of the educational community for baseline data of the type provided by this study. Studies of the future should examine reasons why subjects exhibit the levels of marine awareness they do in other geographic regions, but need not necessarily make comparisons of awareness between geographic strata.

REFERENCES

Fortner, R. W., and T. G. Teates. Baseline studies for Marine Education: Experiences Related to Marine Knowledge and Attitudes." Journal of Environmental Education, 11 (4): 11-19, 1980.

Goodwin, H. L. An Introduction to Marine Education. Newark, D. E., 19711: Delaware Sea Grant College Program, College of Marine Studies, University of Delaware, April 1977.

Goodwin, H. L. A Statement on the Need for Marine and Aquatic Education. Newark, D. E., 19711: Delaware Sea Grant College Program, College of Marine Studies, University of Delaware, March, 1978.

IN RESPONSE TO THE ANALYSIS OF

Fortner, R. W. and T. G. Teates. "Baseline Studies for Marine Education: Experiences Related to Marine Knolwedge and Attitudes." by Richard M. Schlenker. Investigations in Science Education, 9(2): 39-43, 1983.

Rosanne W. Fortner
The Ohio State University

Thank you for the opportunity to comment on Dick Schlenker's review of my "Baseline Studies" article for I.S.E. His critique is a good one, and in response to his comments, I would like to make I.S.E. readers aware that the reliability of the Survey of Oceanic Attitudes and Knowledge ranged from .62 on Form B to .71 on Form A. All forms of the instrument and item analyses are available from me at 2021 Coffey Road, Columbus, Ohio 43210.

LEARNING, COGNITIVE DEVELOPMENT

Linn, Marcia C., Steven Pulos, and Adrienne Gans, "Correlates of Formal Reasoning: Content and Problem Effects." Journal of Research in Science Teaching, 18(5):435-447, 1981.

Descriptors--*Abstract Reasoning; Academic Ability; Cognitive Ability; *Cognitive Development; Cognitive Measurement; *Cognitive Processes; Cognitive Style; Grade 7; Locus of Control; *Measures (Individuals); Science Education; Secondary Education; Secondary School Science; *Student Characteristics

Expanded abstract and analysis prepared especially for I.S.E. by Anton E. Lawson, Arizona State University.

Purpose

This paper addressed content and problem effects in formal reasoning skill. The questions were 1) how differences between task content (e.g., whether the task is about rods bending or springs expanding) influences performance, 2) how differences in task problems (e.g., whether the subject is asked to set up an experiment or to criticize a conclusion) contribute to performance on formal reasoning tasks, and 3) how content and problem effects interact with ability and personality factors previously identified as related to formal reasoning. No hypotheses were specifically stated.

Rationale

Piaget's structural theory of formal thought suggests that a general construct of formal reasoning exists. Emerging in preadolescence, formal thought is consolidated, presumably universally, by adulthood. In recent years, this assertion has come under serious debate. Also formal reasoning, as distinct from previous measures of general ability and achievement, has been questioned. Regardless of these issues, reasoning strategies such as controlling variables, proportions, and functions, have practical validity. Thus improved understanding of the development and use of these strategies remains an important educational issue. The approach taken in the present study was to assume that the general construct of formal reasoning is valid and to scrutinize the current research methodology in hopes of better understanding the reasons behind students' performance on formal tasks. To do this, performance on formal reasoning tasks, which were varied in content and problem selected, was compared to performance on a variety of ability and personality measures.

Research Design and Procedure

The subjects (Ss) were 124 seventh graders from an urban middle school of mixed SES background (61 male, 63 female, X age = 13.46, SD = 0.53). Each S was individually tested on three tasks involving the control of variables schema called Springs, Bending Rods, and Ramp. Order was counter-balanced. Problems raised within each task were called naming variables, controlling, and analysis. For naming variables Ss were awarded a single point for each independent variable they were able to name that might affect the dependent variable in the task. For controlling Ss were awarded one point for each independent variable they successfully tested. For analysis Ss responded to three questions about an uncontrolled test of one of the independent variables. They were awarded one point for each correct answer.

Ss were group tested on the following ability and personality measures: Achievement (as measured by the Stanford Achievement Test), Field Dependence/ Independence (as measured by the Portable Rod and Frame Test and an embedded figures test), Locus of Control, Category Width, Divergent Thinking (as measured by Torrance's order task), Hypothesis Testing (as measured by Levine's concept identification test).

Findings

An analysis of variance revealed a significant effect for content and problem type and a significant interaction. The Ramp was generally the easiest of the tasks followed by Springs and the Bending Rods. Correlations among the formal task problems ranged from .06 to .52. Correlations among ability and personality measures ranged from -.14 to .42. Achievement correlated moderately with Hypothesis Testing (.42) and with Locus of Control (.41) while Hypothesis Testing and Locus of Control correlated moderately with each other (.37). Regression analysis of the ability and personality measures on the formal scores showed the ability and personality measures to be quite uniformly related to the controlling and naming variables problems but strong task effects were apparent for the analysis problem. For each task, the controlling problem was related to achievement, the embedded figure test, and locus of control. The Rod and Frame test was related to two of the three tasks and divergent thinking was related to one of them.

Interpretations

The present findings suggest that content and problem effects deserve scrutiny in research on formal reasoning. Formal reasoning was consistently related to the Rod and Frame test and frequently to tests associated with general ability. These findings place formal reasoning as a mix between traditional school learning and strategy utilization. It appears that the Rod and Frame test taps a fairly pervasive characteristic of formal reasoning for tasks with scientific apparatus, measuring ability to select relevant information. The embedded figures test was also related to the formal measures indicating that the tasks require similar cognitive operations perhaps the selection of appropriate strategies as suggested by Pascual-Leone.

From an educational standpoint these results suggest that generalizing formal reasoning to a new content area is not automatic. Since these strategies are generally taught in science classes using science related tasks, they may not generalize to other areas without additional instruction.

ABSTRACTOR'S ANALYSIS

The present study represents an attempt to gather additional information regarding the nature of formal reasoning by administration of three controlling variables tasks and a battery of ability and personality tests to a group of seventh graders. The novelty of the study is in its selection of measures and statistical analyses. The value of the study is due primarily to the fact that the notion of controlling variables is an extremely important one, yet not enough is presently known about why some students come to use the strategy while others do not. The authors believe, and I concur, that a means of gathering data to help answer this question is to explore individual personality and ability differences in the rates of attainment of the controlling variables schema.

Although the authors fail to explicitly state any hypotheses, one has the impression that a primary motivation for the research was to demonstrate (1) that performance on controlling variables tasks varies from one task to the next (the authors' content effects) and, (2) that performance varies with the type of questions asked within each task (the authors' problem effects). In this sense, the study was successful as content and problem effects were demonstrated. Exactly what the demonstration of problem effects contributes to the literature, however, is not clear. Theoretically seventh graders (13-year-olds) are not supposed to be fully formal operational and at "equilibrium" with respect to formal structures. Thus content effects are theoretically supposed to exist and have been found in many previous studies. During the concrete stage, content is pervasive and general structures for dealing with hypothesis testing have not emerged. Therefore one would expect that certain aspects of performance would vary considerably with familiarity (i.e. content).

The issue of problem effects is an interesting one and essentially unique to this study. The fact that they were found represents a worthwhile contribution to the literature. Consider, for example, the relationship between naming and controlling variables. Anyone who has administered the Bending Rods task to a large number of students is struck by some students' failure to name potentially influential independent variables. It would certainly seem that a variable will not be controlled if it is not first identified. Thus one would expect performance on the naming variables problem to be highly correlated to performance on the controlling variables problem. The relatively low correlations found were somewhat surprising. The authors' explanation for this in terms of field dependence, (i.e., the field dependent students have the knowledge or ability to name variables but do not prefer to) is an interesting one. One might then ask, why do they not prefer to name the variables? No answer to this question is offered.

With regard to the more general issue of individual differences, one set of statistics was not reported that I would have found helpful. The authors did not report an intercorrelation matrix of the controlling variables task scores and the ability and personality measures. From the regression statistics one can see that high achieving, field independent, internal, divergent thinkers do generally better on the controlling variables tasks. Unfortunately only the degree of correlation, after achievement was entered into the regression equation, was shown. The authors state that achievement was entered first because it "is a well understood construct" but I really doubt that this is the case. Achievement is an extremely complex interaction of a number of variables such as interest, aptitude, and opportunity, each of which can in turn be analyzed into subvariables such as field dependence/field independence, locus of control, mental capacity and the like. Given the complexity of the achievement variable, and what I perceive to be the difficulty in identifying just what is involved in achievement, I would have preferred to see the initial correlations among the controlling variables tasks and the ability and personality measures, and I would have preferred the regression to be done with the ability and personality variables entered without inclusion of the achievement variable.

Hypothesis testing ability, as measured by Levine's concept identification test, did not correlate well with formal task performance. This was unexpected to the authors in that the schema of controlling variables most certainly

appears to be involved in the testing of hypotheses. Perhaps the reason for the low correlations is that hypotheses testing on Levine's test involves noncausal situations while the controlling variables schema may be called for only in causal situations (i.e., *in the testing of single causes embedded in* potentially multiple cause-effect situations). Lawson (in press) found hypothesis testing performance to vary considerably between causal and noncausal situations.

The major educational implication of the study was that, for generalization to occur, instruction in controlling variables should not be restricted to science problems. This seems a most reasonable assertion yet the study really did not directly address the issue of instructional generalization. A number of previous training studies have more directly addressed this issue. Training studies, in which aspects of formal reasoning are taught with various methods and materials to students at various ages which are then tested with tasks from various content areas, seem a much better way to determine the degree of generalization which can be expected following different types of instruction.

Science progresses generally via theory construction and the testing of hypotheses generated from theory. The authors' failure to explicitly state testable hypotheses in the present study proved somewhat troublesome. Perhaps this was appropriate given the descriptive nature of the study, yet I believe, if possible, it would be a considerable benefit to the reader if research reports began with clear questions followed by explicitly stated hypotheses and predictions. The study lacked this clear train of thought thus did not tie together conceptually and its relation to theory was left unclear.

Incidentally Tables I and II somehow were reversed in the paper. Also the reader was asked to refer to Table I to note that little variance for naming variables on the Ramp task emerged from the data, yet this statistic was omitted from the appropriate table. These oversights detracted from the paper's readability.

REFERENCES

Lawson, A. E. "The Effects of Causality, Response Alternatives, and Content Continuity on Hypothesis Testing Reasoning," Journal of Research in Science Teaching, Vol. 20 (in press).

Shayer, Michael, Philip Adey, and Hugh Wylam. "Group Tests of Cognitive Development Ideals and a Realization." Journal of Research in Science Teaching, 18(2):157-168, 1981.

Descriptors--*Cognitive Development; *Cognitive Measurement; *Cognitive Tests; *Developmental Stages; Elementary School Science; Elementary Secondary Education; Evaluation Methods; *Group Testing; Intellectual Development; Science Education; Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Edmund A. Marek, The University of Oklahoma.

Purpose

The purpose of this research report is to describe seven Science Reasoning Tasks (1979) beginning with the development of the SRTs through the validation and utilization. Item discriminations, reliability and validity are assiduously reported.

Rationale

Cognitive development has traditionally been determined by the methode clinique developed by Jean Piaget. The investigators identify four essential features of the classical method of measuring levels of cognitive development: 1) allowance for the child to be influenced by his perceptions and the apparatus; 2) opportunity to investigate the reasons for the child's responses; 3) ability to observe the child's reaction to interviewer feedback; and 4) opportunity to question the child's response. Interview methods are very time-consuming, making it impossible or very difficult to collect large quantities of data for this type of research. The researchers developed tests of cognitive development which can be used to assess individuals, in groups of twenty or more, simultaneously.

Research Design and Procedure

Development of Science Reasoning Tasks is summarized by the investigators in five statements:

1. Selecting from the tasks devised by Piaget et al. (1956, 1958, 1964, 1974) those which cover the range of stages to be studied, and which seem most likely to be transposable to a group situation.
2. Writing test items from questions reported by Piaget and Inhelder in their interview tasks, together with appropriate instructions for administration.
3. Ascribing developmental stages to each possible reply to each item, followed Piagetian protocols. In practice, almost all items test the attainment of just one level or sublevel and a complete task must include items covering a suitable range of stages.
4. Devising an overall marking scheme by which a level may be ascribed to a pupil on the basis of his replies to a series of items. In general a two-thirds rule is followed: if there are six items at stage n, then four must be correct to indicate achievement of that stage.
5. Trying the task on a sample of pupils and assessing each by the provisional marking scheme.

Item discriminations and reliability, of the tasks developed by this process, were determined using item discrimination diagrams for tests assessing the range from 2B to 3B. Content validity was established by producing an adequate number of items at each of the levels 2B, 2B/3A, 3A and 3B. Internal consistency was measured by the Kuder-Richardson formula 20 at each stage of development of a task. Reliability was estimated with test-retest correlations tests given three months apart.

Findings

Examination of KR20 correlations indicates that the SRTs are virtually the same as the original Inhelder or Piaget Tasks. The predictive validity was determined with two SRTs -- Tasks II and III. Above average 11-13-year-old students were tested with a 50-60 item content examination in physics, chemistry, and biology. The researchers report that these questions measure understanding rather than recall. Predictive validity correlations of .77 and .78 were measured for the two sections of the course.

Construct validity was established by administering SRT Tasks III-VII to approximately 560 students, age fourteen years. Factor analysis showed a single factor accounting for 59 percent of the total variance. Population surveys with the group tasks have been conducted and population norms established for cognitive development of British school students. A wide range in rates of cognitive development was found and the correlation of Piagetian stages and age was 0.35. Population norms of these Piagetian measures do not increase after the adolescent growth spurt and the researchers predict this could account for sex differentials researchers in the United States have found on formal operations with college students.

The SRTs have also been used in studies in other cultures: southeast Asian countries, West Indies, the Philippines, Palestine, Zimbabwe-Rodesia and Swaziland.

Interpretations

The Science Reasoning Tasks have been thoroughly and carefully developed, validated, and reported. SRTs have much potential in applied research. The researchers conclude that the most powerful use of the tasks may be with matching teaching/learning activities with the cognitive developmental level of the learner. Interpretations from this instrument development research are reported throughout the manuscript and summarized in this statement:

> By monitoring the progress of groups of individuals, whose performance on SRTs has been recorded, through the curriculum and noting areas of success and failure, we can gain real insight into the levels of cognitive development needed to successfully complete each small section of the curriculum. In this way difficulties can be differentiated into those which can be remedied by changes in teaching approach and those which demand restructuring or even complete reframing of the curriculum.

ABSTRACTOR'S ANALYSIS

Development and validation of group tests of cognitive development are important achievements for research in science education. The investigators of this study conducted a comprehensive research and development program to produce the Science Reasoning Tasks. This well written research report thoroughly documents the instrument development and field testing of those instruments. The investigators thoroughly reported each step of the research and included the essential historical influences--i.e., the Piaget interview tasks of cognitive development--providing the chronological evolution of the SRT development.

Group tests of cognitive development should provide data to identify the students' thought as preoperational, concrete operational, formal operational or transitional between levels. The SRT developers state that pupils' development should be categorized on at least six levels of cognitive development:

1	preoperational
2A	early concrete operational
2B	late concrete operational
2B/3A	transitional between late concrete operational and early formal operational
3A	early formal operational
3B	late formal operational

Previous research by Longeot (1965), Warburton (1966) and Raven (1973) used paper-and-pencil tests to measure logical reasoning and categorize thinking as concrete operational or formal operational. Tisher (1971) used a paper-and-pencil test on which the students were required to answer all the questions as thought experiments. Rowell and Hoffman (1975) developed a group test which required a set of apparatus for each student. Recognizing the limitations of these group tasks, the investigators developed a valid and reliable set of demonstration plus paper-and-pencil tasks of cognitive development. The Science Reasoning Tasks were then field tested and refined.

The SRTs utilize a set of apparatus to demonstrate various experiments and a series of questions to which the subjects respond in writing. The abstractor agrees with the developers of the SRT that giving a SRT is more like teaching a lesson than giving a standardized test.

REFERENCES

Inhelder, B., and J. Piaget. The Growth of Logical Thinking. London: Routledge and Kegan Paul, 1958.

Inhelder, B., and J. Piaget. The Early Growth of Logic. London: Routledge and Kegan Paul, 1964.

Longeot, F. "Analyse Statitique de Toirs Tests Genetiques Collectifs." Bulletin de l'Institut National d'Etude du Travail et d'Orientation Professionnelle, 1964.

Piaget, J., and B. Inhelder. The Child's Conception of Space. London: Routledge and Kegan Paul, 1956.

Piaget, J., and B. Inhelder. The Child's Construction of Quantities. London: Routledge and Kegan Paul, 1974.

Raven, R. J. "The Development of a Test of Piaget's Logical Operations." Science Education, 57:377-385, 1973.

Rowell, J. A., and P. J. Hoffmann. "Group Tests for Distinguishing Formal from Concrete Thinkers." Journal of Research in Science Teaching, 12:157-164, 1975.

Science Reasoning Tasks. Seven Piagetian Group-tests and a General Handbook. Windsor: NFER, 1979.

Shayer, M. "The Analysis of Science Curricula for Piagetian Level of Demand." Studies in Science Education, Leeds, 5:115-130, 1978.

Tisher, R. "A Piagetian Questionnaire Applied to Pupils in a Secondary School." Child Development, 42:1633-1636, 1971.

Warburton, F.W. "Construction of the new British Intelligence Scale: Progress Report." Bulletin of the British Psychological Society, 19:68-70, 1966.

Coulter, David, Harvey Williams, and Henry Schulz. "Formal Operational Ability and the Teaching of Science Processes." School Science and Mathematics 81(2):131-138, February, 1981.

Descriptors--Algorithms; Chemistry; *Educational Research; *Learning Theories; Problem Solving; *Science Education; *Science Instruction; Secondary Education; *Teaching Methods; Testing

Expanded abstract and analysis prepared especially for I.S.E. by Donald E. Riechard, Emory University.

Purpose

The stated purpose was to investigate one of the mechanisms that teachers may use to convince themselves that they have taught concepts which Piaget-based theory and research imply cannot be learned by at least 50 percent of their students. Research hypotheses were not specified.

Rationale

An assumption was made that achievement of the objectives of the new post-Sputnik science curricula (IPS, PSSC, BSCS, etc.) requires learners to function at Piaget's formal operational level. The work of Lawson and Renner (1974) and of Karplus and Peterson (1970) cited to support the contention that most high school students do not function at the formal level.

The authors maintain, however, that concrete operational students can be successful in the new curriculum courses because 1) teachers teach algorithms for type-problems, 2) students who are concrete operational memorize the algorithms, and 3) when confronted with type-problems, students are able to apply the appropriate algorithm. Thus, students might get the correct answers but they have not, in fact, used formal operational thinking to do so.

Research Design and Procedure

The sample was composed of 61 ninth and tenth grade students in an urban high school. The subjects were enrolled in an Introductory Physical Science (IPS) course offered as an elective for university preparation. Concrete and formal operational students were identified by a paper and pencil adaptation of the Piaget rod bending activity. Student achievement levels in IPS were determined by average performance on teacher-made tests.

Students' abilities to solve IPS problems consistent with the course's formal operational objectives were evaluated by teacher-made tests and experimental test items. On two occasions during the winter terms, two experimental test items were included in the teacher-made chapter tests administered by the regular classroom teacher. The experimental items were designed so that their solutions could not be derived by merely applying a previously learned algorithm.

Findings

Of the 61 subjects, 14 tested formal operational on the rod bending test. Twelve of the 14 were classified as high achievers and 2 as middle achievers. Twenty two students were in the transitional stage between formal and concrete

operational. Seven of the transitional students were high achievers and 15 were middle achievers. Of the remaining 25 students who tested concrete operational, 6 were high achievers, 4 middle, and 15 were low achievers.

Only a 3.1 percent difference was found between the means of concrete and formal thinkers on the teacher-made tests. The difference was 19 percent on experimental test items. The experimental test items discriminated between high achieving formal and concrete operational students at the .05 level. Teacher-made tests did not.

Interpretations

Piaget's theory was supported. Further, results suggest that teachers do teach algorithms for type-problems. Students who do not understand the concepts are able to meet teacher expectations by memorization of algorithms. Thus, teachers may convince themselves that their students are achieving the objectives defined by the course.

The question was raised as to whether the logic of the learner or the logic of science (as presented in the new curricula) should be the basis for science programs. The authors feel their results provide evidence that the gulf between the developing logic of students and the logic of science is greater than suspected when the post-Sputnik science programs were conceived. They conclude that more emphasis must be placed on the logic of the child.

ABSTRACTOR'S ANALYSIS

Written Report

The written report adequately conveys the general intent, procedure, and results of the investigation. Some changes, however, would improve the readability and sharpness of the paper. The stated purpose, for example, is rather vague. Inclusion of specific research questions or hypotheses would focus the reader's attention on precisely what was to be investigated and give clues to the adequacy of the research design.

Classification of students into achievement levels was not mentioned in the statement of purpose. The abstractor, therefore, was not sure how the achievement variable was to be used or if it was used properly. Vagueness of the purpose also leaves the abstractor unable to comment specifically on the appropriateness of the statistical analysis or presentation of results (tables, etc).

Research Design and Validity

The validity and reliability of any study can be no better than the instrumentation used or the basic assumptions underlying the research design. The rod bending test used here was designed and validated by another researcher whose work is cited. But how valid and reliable was the test? What means were used to determine validity and reliability of the experimental items and the teacher-made tests?

The report suggests that two experimental items were administered on each of two occasions. How does that number (four) compare to the number of teacher-

made items? Do four items adequately sample a student's ability to solve problems requiring formal operational thought? And finally, how valid is the assumption that attainment of the objectives of post-Sputnik curricula, specifically IPS, requires students to think formally? While I tend to agree, I am not aware of a body of research which supports the assumption.

The design of this investigation does not permit generalization beyond the sample studied. The sample is small and seems to have been one of convenience. Sample size alone, however, is not the major concern. The same size sample randomly drawn from a relatively large population involving different schools, teachers, and students would allow generalization to the total population. Further, some of the dangers to validity and reliability would be minimized.

It should be noted that the investigators clearly recognized the hazards of limited sampling and cautioned against drawing sweeping conclusions from their study. The abstractor is also aware that, from a practical point of view, there are many barriers to random sampling from a large multi-school population. The sample of convenience is often the best we can do in a given situation. However, randomization remains one procedure which could strengthen the design of most studies.

Related Research

The finding that less than one third of the students tested were at the formal operational level is consistent with a relatively large body of research. To pursue the topic of formal reasoning in more detail, the reader might want to examine the references and research reported by Hale (1983), Karplus (1977), Lawson (1982), Linn (1982), Piaget (1972), and Renner et al. (1976).

There has been much folklore about the failures of the post-Sputnik curriculum projects. However, a good bit of published research on student performance in those new programs shows a positive effect. A recent study by Shymasky et al. (1982), for example, showed rather impressive gains by students using the new curricula. The IPS students in this study were also successful, as measured by teacher-made tests, in meeting the course objectives.

Contributions and Future Research

The major contribution of this study is not that its findings are consistent with the related research presented above. Rather, its contribution lies in the effort to determine why students appear to be successful in the new science despite their inability to perform the kind of thinking assumed to be required for success.

Persons who have invested heavily (financially and/or emotionally) in the new programs might find it hard to accept the findings and implications of this investigation. However, my guess is that answers to some of the questions posed would tell us a great deal about why students move through our science programs, K-college, and yet so misunderstand the scientific enterprise as to be illiterate in science.

The study raises extremely important questions about curriculum development. It also suggests we have forgotten Bruner's (1960) advice that we should teach in an intellectually honest form. Additional research is needed to corroborate

the findings of this investigation and further examine science education policy as it relates to the needs of students and society.

REFERENCES

Bruner, J. The Process of Education. New York: Vintage Books, 1960. p. 33.

Hale, J. "Problem-Solving Analysis: A Piagetian Study." Journal of Research in Science Teaching 20(1): 77-85, 1983.

Karplus, R. "Science Teaching and the Development of Reasoning." Journal of Research in Science Teaching 14(2):169-175, 1977.

Karplus, R. and R. Peterson. "Intellectual Development beyond Elementary School II." School Science and Mathematics 70:813-820, 1970.

Lawson, A. "The Nature of Advanced Reasoning and Instruction." Journal of Research in Science Teaching 19(9):743-760, 1982.

Lawson, A. and J. Renner. "A Quantitative Analysis of Responses to Piagetian Tasks and its Implication for Curriculum." Science Education 58(4):545-559, 1974.

Linn, M. "Theoretical and Practical Significance of Formal Reasoning." Journal of Research in Science Teaching 19(9):727-742, 1982.

Renner, J., D. Stafford, A. Lawson, J. McKinnon, F. Friot, and D. Kellogg. Research, Teaching, and Learning with the Piaget Model. Norman, OK: University of Oklahoma Press, 1976.

Piaget, J. "Intellectual Evolution from Adolescence to Adulthood." Human Development 15:1-12, 1972.

Holliday, William G. "Selective Attentional Effects of Textbook Study Questions on Student Learning in Science." Journal of Research in Science Teaching, 18 (4):283-289, 1981.

Descriptors--*Attention Control, *Biology, *Diagrams, High School Students, Models, Pictorial Stimuli, Science Education, Science Instruction, Secondary Education, *Secondary School Science, Visual Aids

Expanded abstract and analysis prepared especially for I.S.E. by Edmund A. Marek, The University of Oklahoma

Purpose

The purpose of this research investigation was to test the hypothesis that an adjunct population-question or a no-question treatment would independently be more effective than a sample-question treatment on comprehension. Comprehension was measured with an achievement test.

Rationale

Research investigating the effectiveness of using questions as study aids has been referred to as "mathemagenic" research. Mathemagenic research (Markle and Capie, 1976, and Holliday, 1979) provides evidence of some effect on achievement by using questions as study aids. Furthermore, research by Anderson (1970) and Wittrock and Lumsdaine (1977) suggests that partial-question sets and complete-question sets may have different effects on achievement when compared to a no-question control group. The critical review by Wilson and Koran (1976) of mathemagenic research is one of the few attempts to examine how learners react to questions as study aids. The research abstracted here utilizes the previous work and conditionally generalizes to textbook study questions as they affect learning of a science flow diagram.

Research Design and Procedure

The sample consisted of 176 high school biology students, ranging in age from 14 to 16 years, enrolled in two Calgary (Alberta, Canada) public schools. Students were randomly assigned to one of four groups, three treatment groups and a control group. The three treatment groups were provided with a picture-word diagram describing biogeochemical cycles. One treatment group, the population-question group, received the diagram and twenty textbook study questions. The sample-question group received the same diagram and five textbook study questions. The no-question group received just the diagram and the control group received the science content presented in prose passage form.

The diagram -- colored, picture-word, stylized line drawings -- presented the oxygen, carbon dioxide, nitrogen and water cycles. The study questions developed for this research covered all critical information associated with a given concept displayed in the diagram. A complete description of the diagram cycles and the study questions can be found in this article. Students were instructed to self-pace themselves in learning the presented material. They were encouraged to take plenty of learning time. All of the students were tested at the same time with "a 30-question multiple-choice test containing only words."

A post test-only equivalent-group design measured comprehension with an achievement test. This instrument covered the same diagram content of the 20 study questions, although it contained no diagrams, and was based on content synthesis of two or more of the study questions.

Findings

Analysis of variance of the four group scores from the post test were significant at $p < 0.05$. The Student Newman-Keuls multiple range procedure supports all hypothesized group orderings ($p < 0.05$): population-question and no-question, sample question, and control. Internal consistency or reliability using Cronbach's alpha was 0.83 for the achievement test.

Interpretations

Encouraging or aiding students with selectively focusing their attention on a sample of criterial content can result in reduced comprehension of such specialized instructional materials as flow diagrams. The selective attention model of this study also predicts a similar inhibitory learning effect occurs when prose passage is used to present flow diagrams. The results of this study show that students, provided with a complete study aid (population-question treatment) or no study aid (no-question treatment) while covering science content in a flow diagram, can achieve greater comprehension than students provided with a partial study aid (sample-question treatment). Instructional systems aimed at focusing learners' attention on a sample of the criterial content can reduce comprehension or learning effectiveness.

From this research investigation Holliday concludes that this selective attention model is a simple and reliable basis for evaluating learning hypotheses in science education, yet it is seldom used.

ABSTRACTOR'S ANALYSIS

The textbook is the common denominator of science programs in this country and perhaps throughout the world. Research of textbook usage, including reading chapters and answering study questions, would seem to be a logical and important research effort. Mathemagenic research is an example of a matrix of research resulting from a vigorous quest into textbook usage as the focus of science programs. In this regard the research abstracted here has contributed to science education because the research investigation offers data about a technique frequently dominating science programs. That technique is the utilization of textbook study questions as study aids.

The researcher designed an investigation to analyze the effectiveness of using study questions to assist students with learning selected science context -- biogeochemical cycles presented in a flow chart. The merit of the selection of this research problem comes from the frequency of the practice of using textbook study questions to facilitate learning. The abstractor offers alternative or additional foci when investigating comprehension in science: the nature of science and the cognitive developmental level of the learner.

Dr. Duane Roller, a historian of science, defines science as the quest for knowledge, not the knowledge. Science involves more than the scientific information; science is the acquisition of the scientific information. The

acquisition process includes more than reading about scientific precepts, discoveries and technological advances. Science is measuring, observing, experimenting, model building, interpreting, predicting. Research in science education should presuppose this dimension of science.

Perhaps research in science education should treat the textbook as a supplementary component of the science classroom and not the primary component. Therefore, extensive research matricies, such as mathemagenic research, will assume a secondary status and science education research will deal with the discipline of science as science and not reading. The abstractor's analysis is not suggesting that the research defined science as only reading.

This research project investigated the learning of biogeochemical cycles ordinarily taught to the tenth-grade students in two Calgary (Canada) public schools. A mismatch exists with the content and the ability of the learner to assimilate the content. Shayer and Adey (1981) classified biogeochemical cycles as early formal operational (3A). The results of the Holliday investigation may be affected by this mismatch if the students of the study (14-16 year olds) are not early formal operational. Mathemagenic research must consider this mismatch if empirical evaluations of printed materials and teacher questioning behaviors are to continue.

The research design, statistical analyses and written report were well organized and succinctly reported. However, a suggestion for improvement is offered for the bibliographic references. Numerous inconsistencies exist in the references of this manuscript published in the Journal of Research in Science Teaching. One article is cited with quotations while the remaining are cited without quotations. Journals and books are referenced with all lower case letters, except, of course, the first word of the title, for some citations and not for others. The research by Anderson (1970) cited in the introduction of the journal article is not included in the references or else the date is inaccurately recorded.

REFERENCES

Anderson, R. C. "How to Construct Achievement Tests to Assess Comprehension." Review of Educational Research, 42:145-170, 1972.

Holliday, W. G. "How to Write Effective Study Questions." A NARST paper presented at the National Science Teachers Association, Atlanta, March, 1979.

Markle, G., and W. Capie. "The Effect of the Position of Inserted Questions on Learning From an Activity-Centered Science Module." Journal of Research in Science Teaching, 1976, 13, 167-170.

Shayer, M., and P. Adey. Towards a Science of Science Teaching. London: Heinemann Educational Books, 1981.

Wilson, J. J., and J. J. Koran, Jr. "Review of Research on Mathemagenic Behavior: Implications for Teaching and Learning Science." Science Education, 60:391-400, 1976.

Wittrock, M. C., and A. A. Lumsdaine. "Instructional Psychology," In M. R. Rosenzueig and L. W. Porter (Eds.) Annual Review of Psychology, Palo Alto, CA: Annual Review Inc., Vol. 28, 1977.

Boulanger, David F. "Ability and Science Learning: A Quantitative Synthesis." Journal of Research in Science Teaching, 18(2):113-121, 1981.

Descriptors--*Ability; *Achievement Rating; *Correlation; Elementary School Science; Elementary Secondary Education; *Learning; Science Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by David L. Dunlop, University of Pittsburgh at Johnstown.

Purpose

The purpose of this article was to review, analyze, and synthesize published studies relating ability measures to science learning measures in order to establish the best estimates of such correlations under various study conditions. No explicit hypotheses were being tested; however, a general question, "What influence does one's ability have on his/her learning of science?", was summarized and discussed.

Rationale

The rationale for this study is based upon the author's belief that the unique features of science instruction such as laboratories, the use of quantitative skills, and the cumulative nature of the subject matter suggest that estimates of correlations of the constructs with general learning outcomes may not be accurate for science learning. Furthermore, Yager (1978) has identified broadly based reviews which draw on general education research as a national need providing that they inform and augment science education research.

Research Design and Procedure

A search of published literature in science education encompassing the period of 1963-1978 was conducted to identify studies which reported a correlation between ability and learning. The search was further limited to studies conducted with subjects in grades 6-12. Ability was defined as any cognitive measure that predicts science learning, and studies including ability measures as blocking variables or as covariates in ANOVA were excluded, unless a zero-order correlation was reported between ability and outcome measures.

The search resulted in the identification of 34 studies which, collectively, contained 207 raw correlations. To insure the independence of each correlation in a given ability-outcome cell, each study's median correlation for a given cell was computed, reducing to 67 the number of correlations used in the analysis.

All assembled studies were numerically coded according to approximately 40 variables, and an independent check by a second researcher of the reliability of coding revealed about 90 percent agreement. The coding process yielded three ability categories and four learning outcome categories, forming a 3 X 4 outcome matrix. The three ability categories were general ability, prior achievement, and quantitative-spacial reasoning. The four outcome categories were factual, product, process, and attitudinal learning.

An overall ability-attitude mean correlation was calculated from five of the correlations which related ability to attitudes. Because of the small number of correlations, no further analysis of the ability-attitude relationship was attempted.

The three cognitive outcomes were analyzed together and were represented by 62 median correlations. A two-way analysis of variance was conducted to determine if the differences among ability categories were attributable to chance. The 62 median correlations were combined across all cells for the best overall estimate of the ability-cognitive outcome correlation, and an analysis of the influence of other variables such as sample size, subject matter, and study design was conducted on the entire 62-correlation data set.

To determine if any study variables systematically biased the reported ability-outcome correlations, the values of study variables were dichotomized into approximately equal subgroups and the t-test was applied to compare each resulting subgroup pair. In addition, correlations of continuous study variables with associated ability-cognitive outcome correlations were computed.

Findings

The overall ability-attitude mean correlation was reported to be .17 with a standard deviation of .07. Since there were only five raw correlations in this category, no additional analysis was conducted on this sub-group.

Correlations between the ability measures (general ability, prior achievement, and quantitative-spatial) and the outcome measures (factual, product, and process) were presented. No correlations existed between quantitative-spatial and process outcomes; hence, that cell was empty. The range of the mean ability-outcome correlations across the eight remaining cells was .41 to .53.

A two-way analysis of variance was conducted to determine if the differences among categories were attributable to chance. Main effects (ability: $F = .46$, $p = .64$; Outcome: $F = .38$, $p = .69$) and interactions ($F = .11$, $p = .95$) were non-significant. When the 62 correlations across all cells were combined, the best overall estimate of the mean ability-cognitive outcome correlation was .48 with a standard deviation of .15.

A t-test was used to compare the two subgroups resulting from the dichotomizing of the study variables. Only one difference (reliability of outcome measure) was significant at the p .05 level.

Correlations of continuous study variables with associated ability-cognitive outcomes resulted in only one significant (p .05) correlation, and that was the reliability of the outcome measure.

Interpretations

The tenet that ability and past learning are among the best predictors of future learning is well established among educational researchers and practitioners. However, it is not totally clear how this tenet applies to specific situations with varying subject content and study conditions. The estimates developed in this study should provide the researcher in science education with a guide for estimating the influence of ability on science

learning in an untested population, as well as a norm for comparing new findings on the extent to which various factors influence learning.

The summaries provided by this study highlight the fact that measured ability, on average, does not account for a great amount of variance in science learning. Specifically, the five ability-attitude outcome correlations gave a mean of .27 with a standard deviation of .07, while the 62 ability-cognitive outcome correlations had a mean of .48 with a standard deviation of .15. Although ability predicts cognitive outcomes better than attitudinal outcomes, neither accounts for a major portion of the variance.

Several other factors such as student motivation, quality and quantity of instruction, home environment, etc. are known to affect learning. As improved estimates of the effects of these other factors on science learning become available, science education research and teaching practice can be directed at optimizing those influences most potent in improving science learning, keeping the less manipulative ability factor in proper perspective.

ABSTRACTOR'S ANALYSIS

This study reviews the correlational data concerning ability and the learning of science, and it is important for at least two reasons. First, it demonstrates that, on average, ability accounts for approximately 23 percent of the variance in science learning. This leaves a significant amount of uncertainty in the prediction of science learning and should stimulate additional research. Second, there is a great need to periodically review, analyze, and synthesize the published data in a given discipline or line of research.

Although generally well done, the article was unclear as to the rationale for dichotomizing the values of the study variables into equal subgroups. This presumably means an equal number of correlations; however, one could assume an equal number of students. In either case, what is the rationale for an equal number? The criteria for selecting cut-off points should be defined and have a rationale which relates to the characteristics of the given subgroup. A "top half" versus a "bottom half" comparison may result in different findings than a "top third" versus "bottom two-thirds" comparison.

Another point of minor confusion relates to the author's point concerning the relationship between science learning and learning in a broader context. The author states that the unique features of science instruction suggest that estimates of correlations of the constructs with general learning outcomes may not be accurate for science learning. Although this may be true, no evidence was presented in the article to support or refute this point.

The correlations reported in this study are similar to several correlations between IQ and academic achievement as summarized by Jensen (1981). The range of correlations listed by Jensen is .30-.84. He states that the validity of IQ for predicting academic achievement decreases at higher levels of schooling. A typical range for elementary school children would be .60-.70, while a typical range for graduate school students would be .30-.40.

In an investigation of this type where several different sets of data are being synthesized, it is very important that the constructs of each study are clearly identified and properly categorized and/or grouped. This appears to

have been done in this case. For example, a working definition of ability was established, and all studies were numerically coded according to a specified list of variables. Further, an independent check of the coding was conducted by a second researcher.

When attempting to predict an outcome, you must first segregate the relevant variables; however, the variables involved in a given learning situation are very difficult to identify. In fact, we do not know how many are involved nor to what degree they are important. One approach would be to identify a single variable and examine its effects in detail. To an extent this is the approach taken in this study which has focused on student ability as a relevant variable in the learning process. Although several other variables are mentioned, additional factors need to be considered before we can state with an acceptable degree of certainty that ability accounts for an average of 23 percent of the variance in science learning. (Although this may be true for the studies reviewed in this article.) For example, concrete and formal students, as defined by Piaget (1958), may respond differently in a specific learning environment. If this variable (level of cognitive growth) is not considered, the results of the study could be skewed in an undetermined manner. DeCecco (1968) identifies several temporary states of the learner as still another set of variables which could have an effect on learning. These would include conditions such as fatigue, habituation, or drugs. Again, these factors would be very difficult to control in an actual experiment, but could significantly impact the results of the study. Dunlop and Fazio (1976, 1977) have identified student preference as an important variable in the manner in which students choose to solve problems. It is likely that this same variable is important in the learning process as well.

Jensen (1981) reviewed the literature on the relationship between learning and IQ and listed several generalizations which could be of interest to the researcher interested in the factors which govern learning. His list included the following conditions which increased the correlation between learning and IQ:

1. When learning is intentional
2. When the material to be learned is hierarchical
3. When the material to be learned is meaningful
4. When the learning task permits transfer of knowledge
5. When learning is insightful
6. When the material to be learned is of moderate difficulty and complexity
7. When the amount of time for learning a given amount of material is fixed for all students
8. When the learning material is age-related
9. When the learners are at an early stage of learning something new

It remains for an ambitious researcher (or coordinated team of researchers) to make an exhaustive analysis of the different "types" of science learning and then to identify and evaluate the impact of all relevant variables. In the meantime, studies such as this one must continue to stimulate new research and add to the existing body of knowledge through careful review, analysis, and synthesis of published reports.

REFERENCES

DeCecco, John P. The Psychology of Learning and Instruction. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

Dunlop, David L. and Frank Fazio. "Piagetian Theory and Abstract Preferences of College Science Students." Journal of College Science Teaching, 5:297-300, 1976.

Dunlop, David L. and Frank Fazio. "Piagetian Theory and Abstract Preferences of Secondary Science Students." School Science and Mathematics, 675:21-26, 1977.

Inhelder, B. and J. Piaget The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Jensen, Arthur R. Straight Talk About Mental Tests. New York: The Free Press, 1981.

Boulanger, F. David and Barbara K. Kremer. "Age and Developmental Levels as Antecedents of Science Learning." Journal of Research in Science Teaching, 18(4):371-384, 1981.

Descriptors--*Academic Achievement; *Cognitive Development; Elementary School Science; Elementary-Secondary Education; *Predictor Variables; Science Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Rosemary K. Lund Padilla, University of Georgia

Purpose

The purpose of this investigation was to quantitatively synthesize studies conducted between 1963 through 1978 which related age (or grade) and developmental level to science learning among 6th through 12th grade students.

Rationale

Over the past 20 years the developmental perspective has stongly influenced investigations in science education. In particular, Piagetian-based studies relating level of cognitive development and age to various aspects of science achievement were common. However, the strength of the relationship between age and/or developmental level and science achievement was often assumed. This assumption was often based on qualitative analyses provided by various articles and publications such as the annual Summary of Research in Science Education. Even excellent qualitative analyses typically described only possible inter-relationships among the various findings of independent investigations. Until the present study, no quantitative synthesis of the findings of related investigations was attempted. A quantitative synthesis has the advantage over qualitative analyses of being more objective. With this objectivity it is possible to combine and compare findings relating different independent variables, such as age or developmental level to a common dependent variable, such as science achievement.

Research Design and Procedure

The quantitative synthesis used was that advocated by Glass (1978) which is commonly referred to in the literature as meta-analysis. The separate investigation results of interest were either correlations or effect sizes.

When two grade levels, or two age levels, were compared. Boulanger and Kremer calculated an effect size (ES) using one of two formulas:

$$ES = \frac{X_H - X_L}{S_H} \quad \text{or} \quad ES = t \sqrt{\frac{1}{n_H} + \frac{1}{n_L}}$$

where

X_H and X_L are the dependent variable means of the higher and lower grades, respectively;
S_H is the standard deviation of the higher grade scores
t is the t-test statistic
n is the group sizes

The literature search, covering 1963 through 1978, included:

1. all ERIC annual reviews of science education research;

2. a manual search of the 1963-1978 volumes of Journal of Research in Science Teaching and Science Education, and;

3. a computer search of Dissertation Abstracts and Social Sciences Citation Index.

Two kinds of grade 6-12 investigations in science education were sought: 1) those that reported a correlation of age, grade or developmental level to some measure of science achievement and/or 2) those that reported measures of developmental level or science achievement at two different grade levels such that an effect size (ES) could be calculated. Twenty-seven investigations were selected.

Boulanger and Kremer defined developmental level as "any measure of Piagetian stage or related logical operation . . . or other measure validated against Piagetian theory." Cognitive achievement was defined as "any measure of factual and/or conceptual learning of science content." Science process learning "was restricted to scores on the Science Process Inventory (Welch and Pella, 1968)."

The 27 selected investigations were numerically coded (with 90% inter-rater agreement) according to the following variables: "1) type, source, and reliability of independent and dependent measures; 2) grade level, ability level, and science subject; 3) ethnic, urban-rural, SES character of the community; 3) study design, methodological flaws, and unit of analysis; and 4) reported correlations or calculated effect sizes."

Findings

Of the 27 studies selected, 7 correlated developmental level with cognitive achievement. Mean correlations rise from 0.28 in grade seven to 0.63 in grade 9, and fall to 0.32 in grade 12. The grand mean is 0.40 (S.D. = 0.14). Of the 27 studies selected, 6 correlated age or grade with developmental level or cognitive achievement. The correlations reported range from 0.00 when based only on the ages of grade 11 science students to 0.57 when based on the ages of grade 4-9 students. Of the 27 studies selected, 15 reported grade level comparisons of developmental level, cognitive achievement and science processes. Effect sizes were calculated from these grade level comparisons. The effect size increases an average of 0.36 each year for developmental level, 0.28 for cognitive achievement, and 0.43 for science process learning. Boulanger and Kremer note that "expressed in percentiles, the (annual) increments indicate the . . . advance of the mean class scores each year from the previous year's 50 percentile point." "Average yearly percentile increases are 14 points for developmental level, 11 points for cognitive achievement, and 17 points for science process learning."

Interpretations

The reported correlations of age or grade with developmental level range from 0.00 to 0.57. This result emphasizes "the inappropriateness of strongly relating age or grade with levels of intellectual development. . ." Within

class variation might be greater than between class variation. The authors state that age or grade are thus poor predictors of intellectual development or ability. Boulanger and Kremer strengthen this point by comparing the grand mean correlation of developmental level and cognitive achievement (.40) to the mean correlation of ability measures with cognitive achievement (.48) found in a previous meta-analysis (Boulanger, 1980b). They found the .48 correlation to be significantly higher (p 0.01) than the .40 correlation.

In addition Boulanger and Kremer examined the possibility of systematic influences of several factors across studies. Among these factors were the reported reliabilities of the outcome measures and the threats to validity due to study design, such as convenience sampling, and use of cross sectional data."

Finally, the effects of training on science achievement (Boulanger, 1980b) are related to the annual mean percentile gain in developmental level. The annual mean percentile gain was 14 percentile points without training at the grade 4-9 interval contrasted with a 30 percentile point gain with training. Boulanger and Kremer conclude that "studies strongly suggest the annual . . . (percentile point) . . . increments in such developmentally related tracts . . . can be increased with appropriate instruction."

ABSTRACTOR'S ANALYSIS

Examination of the nature of the relationship of age and developmental level with science achievement has been the focus of much research in science education. Does developmental level influence what a middle school age student can learn? Is the science curriculum at this age level written for a student only capable of formal operations? What happens to the student not developmentally "ready" who is still only capable of concrete operations? The present study provides some important clues to answer the above questions. Indeed, the authors rightly note that all of their findings and interpretations should be considered as "hypotheses" for further investigation.

The major importance of this study lies in its use of the technique called meta-analysis. By objectively synthesizing the results of previous studies, the authors used a broad data base with the statistical power to provide a unique perspective to the controversy surrounding the search for the "best predictor" of science achievement: age developmental level or student ability.

Before commenting on specific findings of this study, a brief commentary on the meta-analysis technique seems in order. Meta-analysis, or simply an analysis of analyses, was developed by Glass (1978) as a statistical method to integrate results from independent studies. Two techniques of meta-analyses include calculation of treatment "effect size" and calculation of mean correlation values. The former is generally found in experimental studies examining the effect of a treatment, whereas the latter is found in correlational studies. All of the research design and statistical problems inherent in such studies, e.g., sampling problems, restriction of range, etc., are present in meta-analyses. However, because of the much larger numbers of subjects involved in meta-analyses, the resultant findings can be considered more robust than findings generated by single studies. Glass notes that many weak studies can add up to a strong conclusion (Glass, 1978, p. 104).

Boulanger and Kremer's application of meta-analysis is generally sound. Their search for studies, however, might be lightly criticized for covering only 1963-1978: the more years spanned, the better. They might also be criticized for only reviewing the individual volumes of The Journal of Research in Science Teaching and Science Education. Surely additional journals, such as the American Educational Research Journal, the Educational Researcher, or the American Psychologist might have been helpful. Their inclusion of a search through Dissertation Abstracts and Social Sciences Citation Index is commendable in light of recently reported contradictory findings between published and unpublished research (Smith, 1980a).

A major finding of this meta-analysis: that measures of student ability were better predictors of science achievement than were developmental measures requires some comment. The grand mean correlation of ability with science achievement from a previous study (Boulanger 1980a) was .48 (S.D. 0.15). In contrast, the grand mean correlation of developmental with science achievement in the present study was .40 (S.D. 0.14). Although the authors note that they found a significant difference (p 0.01) between these two grand mean correlations, in fact the ability measure still only accounts for 23% of the variability in science achievement. Thus, 77% of the variability is still unaccounted for, and the search for a "best" predictor of science achievement must continue. Boulanger and Kremer did not note this point outright, but tacitly acknowledge it in their concluding remarks in which they state "in order to sort out the unique contribution of each kind of measure (i.e. developmental versus ability) . . . to the prediction of science learning both . . . (kinds) . . . should be administered and later related to . . . achievement and developmental growth." They further note the need for longitudinal studies.

Another finding of this study: that developmental level and cognitive achievement are each linearly related to grade level and gradually increase across grades 4-9 is at times unclear. Although the authors provided very helpful figural representations of these relationships, this reviewer at times lost sight of this important finding due to the authors' lack of clarity. Inclusion of the calculation methods to arrive at incremental effect sizes might better have been in a note rather than in the body of the paper. Notwithstanding, the linear relationship of annual developmental effect size increments across grades 4-9 is an important empirical validation of Piagetian developmental theory. Perhaps the authros could have expanded this, commenting more specifically on the transition phase between concrete and formal operation characterized by this age student.

The finding that the "correlations of age or grade with developmental level . . . emphasize the inappropriateness of strongly associating age or grade levels with intellectual development or ability to use logical operations" deserves comment. "Intellectual development" may be an overstatement here. Only Piagetian-related measures of intellectual development were sampled in this study. Other means of evaluating "intellectual development" might indeed by related to age or grade level.

In summary, this study adds a significant perspective to the literature surrounding the relationship of age and developmental level to science achievement. It also adds significantly to the empirical search for the "best" predictors of science achievement. The use of the research technique

called meta-analysis is indeed the strength of this study. In this vein the only methodological imperfections stem from restricting the years included and the limitation of journals sampled. The major finding that ability is a better predictor than developmental level of science achievement needs to be viewed in light of the amount of variability (77%) in science achievement still unaccounted for. Lastly, all investigators of the relationship of science achievement to ability or to developmental level need to consider this study in planning future investigations. A concerted attempt to collect both ability and developmental data will add significantly to future synthesis attempts to sort out each measure's unique contribution to science achievement.

REFERENCES

Glass, G.V. "Integrating Findings. The Meta-analysis of Research." In L.S. Shulman (Ed.) Review of Research in Education (Vol. 6). Itasca, Illinois: E.E. Peacock, 1978.

Glass, G.V. "Meta-analysis: An Approach to the Synthesis of Research Results." Journal of Research in Science Teaching, 19:93-112, 1982.